Ciência de Dados na Nutrição
Felipe Teodoro Emanuela Araújo João Vitor Lucas Wall Mateo Wall Victor Vanazzi

Table of contents

1	Intr	rodução	2		
	1.1	Contexto	2		
	1.2	O Problema a ser Resolvido	2		
	1.3	Objetivos do Projeto	3		
2	Me	todologia	4		
	2.1	Tecnologias e Ferramentas	5		
3	Des	Desenvolvimento			
	3.1	Arquitetura da Solução	7		
	3.2	Fluxo de uma Requisição (Exemplo: "Despensa Virtual"):	7		
4	Res	sultados	9		
	4.1	Resultados da Análise de Concorrência	9		
		4.1.1 Metodologia da Análise	9		
		4.1.2 Aplicativos Analisados	9		
		4.1.3 Principais Oportunidades Identificadas	10		
		4.1.4 Posicionamento Estratégico	10		
		4.1.5 Comparação de Aplicativos Concorrentes	10		
		4.1.6 1. MyFitnessPal	10		
		4.1.7 2. Tecnonutri	11		
		4.1.8 3. Programa Mais Saúde (Vitat)	11		
		4.1.9 4. Alimente-se	12		
		4.1.10 5. Fastic	12		
		4.1.11 6. MyDietCoach	13		
		4.1.12 7. Diet: WeightLoss	13		
		4.1.13 8. FitLab	13		
		4.1.14 9. FoodAI	14		
5	Cor	nclusão	15		
	5.1	Dificuldades Enfrentadas	15		
	5.2	Próximos Passos e Sugestões	16		
6	Ref	erências	17		

Introdução

1.1 Contexto

O Brasil enfrenta um cenário nutricional complexo e paradoxal, caracterizado pela dupla carga da má nutrição. Por um lado, persistem desafios relacionados à desnutrição e deficiências de micronutrientes em grupos vulneráveis. Por outro, dados de inquéritos nacionais como a Pesquisa Nacional de Saúde (PNS) e a Pesquisa de Orçamentos Familiares (POF) revelam um avanço alarmante do sobrepeso e da obesidade em todas as faixas etárias e classes sociais, consolidando uma epidemia de doenças crônicas não transmissíveis (DCNT). Este panorama contrasta diretamente com as diretrizes do Guia Alimentar para a População Brasileira, um documento de referência mundial que preconiza uma alimentação baseada em alimentos in natura e minimamente processados. A realidade, no entanto, é marcada por um consumo crescente de alimentos ultraprocessados, impulsionado por fatores como praticidade, baixo custo aparente e marketing agressivo, criando um abismo entre a recomendação e a prática diária da população.

1.2 O Problema a ser Resolvido

A principal barreira para uma alimentação saudável no Brasil, especialmente para a população de baixa renda, não é meramente a falta de informação, mas sim a viabilidade e a percepção. Para uma vasta parcela da sociedade, o termo "dieta" é carregado de conotações negativas, sendo associado a punição, restrição, alto custo e pratos sem sabor. O problema se intensifica em "momentos-gatilho" — como o cansaço ao final de um longo dia de trabalho, o estresse financeiro no fim do mês ou a necessidade de uma recompensa imediata na sexta à noite. Nesses momentos, a escolha por fast food, delivery ou ultraprocessados torna-se o caminho de menor resistência, não por falta de desejo por uma vida mais

saudável, mas pela ausência de alternativas que sejam percebidas como igualmente práticas, acessíveis, prazerosas e culturalmente relevantes. O desafio central, portanto, é quebrar o ciclo em que "comer bem" é visto como um privilégio e "comer o que é fácil" é a única opção realista.

1.3 Objetivos do Projeto

Diante deste complexo problema, o objetivo geral deste projeto é desenvolver uma solução de Inteligência Artificial que atue como um "assistente nutricional" prático, empático e acessível, com foco na população de baixa renda. A meta não é prescrever dietas restritivas, mas sim empoderar os usuários a fazerem pequenas e sustentáveis melhorias em sua alimentação diária. Para alcançar este objetivo macro, o projeto se baseia nos seguintes pilares:

- 1. Desmistificar a Alimentação Saudável: Utilizar a IA para reformular a narrativa em torno da "dieta", focando no prazer, na abundância e na valorização da culinária brasileira (como o arroz com feijão). O sistema deve comunicar benefícios de forma positiva e tangível, evitando linguagem punitiva ou excessivamente técnica.
- 2. Oferecer "Trocas Inteligentes" Contextuais: Criar um modelo de busca por similaridade que vá além dos nutrientes. O sistema deverá compreender a intenção e o contexto do usuário (ex: "quero um lanche rápido e gostoso") e sugerir alternativas que sejam não apenas mais saudáveis, mas também realistas em termos de custo, tempo de preparo e disponibilidade de ingredientes.
- 3. Integrar Custo e Acessibilidade como Fatores Centrais: Tornar os dados de preços de alimentos em tempo real um componente fundamental do sistema de recomendação. O objetivo é garantir que as sugestões sejam financeiramente viáveis e ajudem o usuário a otimizar seu orçamento.
- 4. Criar uma Experiência Empática e Orientada à Ação: Desenvolver uma interface (UI/UX) e uma linguagem que sejam acolhedoras, especialmente nos "momentos-gatilho". A solução deve ser proativa, oferecendo ajuda e alternativas fáceis exatamente quando o usuário mais precisa, transformando o que seria um momento de escolha automática e prejudicial em uma oportunidade de decisão consciente e positiva.

Metodologia

Para garantir a agilidade, a adaptabilidade e o foco no impacto social, o desenvolvimento deste projeto será norteado por uma metodologia híbrida, combinando princípios do Scrum (Breves reuniões par alinhamento de tarefas), para gestão de tarefas e entregas iterativas, com o CRISP-DM, para guiar as etapas do ciclo de vida da ciência de dados. Essa abordagem nos permite construir e validar hipóteses rapidamente em sprints de uma ou duas semanas, enquanto mantemos uma estrutura robusta para a exploração de dados e modelagem.

O fluxo de desenvolvimento é dividido nas seguintes etapas macro:

- Coleta e Preparação de Dados: A base do nosso sistema. Esta etapa envolve o web scraping massivo de fontes de dados brasileiras (portais de receitas, blogs de culinária popular) para construir um dataset rico. Em paralelo, utilizaremos a estratégia do "Funil Inteligente": um conjunto de scripts em Python com heurísticas para etiquetar automaticamente cada receita com features de contexto, como score_de_saudabilidade, score_de_complexidade, score_de_custo e tags_contextuais.
- Modelagem e Treinamento da IA: O núcleo do projeto. Utilizando o dataset de trios (âncora, positivo, negativo) gerado na etapa anterior, faremos o fine-tuning de um modelo Transformer (BERT). O objetivo é especializar o modelo para entender o contexto nutricional, financeiro e emocional do nosso público-alvo, gerando embeddings (vetores) contextuais para cada receita.
- Desenvolvimento do Backend e API: Construção de uma API robusta em FastAPI que servirá como o cérebro do sistema, expondo endpoints para as funcionalidades principais, incluindo a busca por similaridade contextual e a lógica de recomendação.
- Desenvolvimento do Frontend: Criação da interface do usuário, com foco em uma experiência (UX) empática, acessível e orientada à ação. A inter-

face será projetada para simplificar a jornada do usuário, especialmente em "momentos-gatilho", como o "Modo Sextou".

- Automação e Orquestração: Implementação de pipelines automatizados para tarefas recorrentes, como a atualização de dados de preços ou a reexecução do funil de geração de triplets.
- Integração Contínua e Entrega Contínua (CI/CD): Adoção de práticas de CI/CD para garantir a qualidade do código, a automação dos testes e a agilidade nas implantações em produção.

2.1 Tecnologias e Ferramentas

- Python como linguagem principal para ciência de dados e backend.
- Algum tipo de banco de dados relacional para a relação dos preços atualizados em tempo real
- BERT (via Hugging Face Transformers) & PyTorch para o treinamento do modelo de linguagem, permitindo a criação dos embeddings contextuais customizados.
- FAISS (Facebook AI Similarity Search) será a base de dados vetorizada.
 Após o fine-tuning do BERT, os embeddings de todas as receitas serão
 armazenados e indexados no FAISS para permitir buscas por similaridade
 em tempo real com baixíssima latência.

Desenvolvimento

O processo de implementação do projeto foi estruturado em fases sequenciais e iterativas. Cada fase foi projetada para construir uma parte fundamental da solução, permitindo validações contínuas e garantindo que o produto final esteja alinhado com os objetivos de impacto social e viabilidade técnica.

- 1. Fundamentação e Coleta de Dados: A fase inicial consiste na criação de um dataset proprietário através de web scraping em portais de receitas brasileiros. O foco é capturar não apenas ingredientes e modo de preparo, mas metadados contextuais ricos, como tags, tempo de preparo e avaliações de usuários.
- 2. Engenharia de Features e Geração dos Triplets: Utilizando o dataset bruto, implementar o script do "Funil Inteligente" em Python com Pandas para gerar scores heurísticos (saudabilidade, complexidade, custo) e, com base neles, gerar automaticamente o dataset de treinamento triplets.csv.
- 3. Fine-Tuning e Geração de Embeddings: Nesta fase, o modelo Transformer neuralmind/bert-base-portuguese-cased será especializado através de um processo de fine-tuning com a estratégia TripletLoss, aprendendo a mapear o texto das receitas para um espaço vetorial onde a proximidade matemática reflete a similaridade contextual.
- 4. Construção da Infraestrutura de Busca e API: Os embeddings gerados seriam indexados em uma base de dados vetorial FAISS. Em paralelo, foi desenvolvida a API principal em FastAPI.
- 5. Implementação do Frontend e da Lógica de Conversação: A interface seria desenvolvida com foco na experiência do usuário. Implementar algum fluxo para orquestrar a comunicação entre a busca no FAISS e a geração de texto da OpenAI.
- 6. Automação e DevOps: Configurados pipelines em GitHub Actions para

CI/CD e workflows no N8N para tarefas recorrentes, como a atualização de preços.

3.1 Arquitetura da Solução

A arquitetura foi desenhada para ser modular, escalável e de alta performance, separando claramente as responsabilidades de cada componente. O sistema opera em quatro camadas principais que se comunicam via APIs:

- Camada de Apresentação (Frontend): É a interface com a qual o usuário interage. Ela é responsável por apresentar as funcionalidades como a "Despensa Virtual" e as "Trocas Inteligentes" e por comunicar-se com a camada de backend.
- Camada de Inteligência (Backend API): O cérebro da aplicação, construído em FastAPI. Esta camada recebe as requisições do frontend, orquestra a lógica de negócio, utiliza o modelo BERT treinado para processar as buscas, consulta as bases de dados e formula as respostas.
- Camada de Dados: Composta por múltiplos sistemas de armazenamento:
- Índice FAISS: Armazena os embeddings vetoriais de todas as receitas, permitindo buscas por similaridade em altíssima velocidade.
- Banco de Dados (Relacional/NoSQL): Guarda os dados estruturados das receitas, ingredientes, preços e informações dos usuários.
- Camada Externa (API da OpenAI): Utilizada para gerar as respostas de texto de forma natural e empática, baseada no contexto fornecido pela nossa API.
- Camada de Automação: Orquestrada por ferramentas como N8N (para scraping de preços e automação de dados) e GitHub Actions (para integração e entrega contínua - CI/CD), garantindo que o sistema seja resiliente, atualizado e facilmente sustentável.

3.2 Fluxo de uma Requisição (Exemplo: "Despensa Virtual"):

O processo de uma interação típica do usuário demonstra como as camadas se comunicam:

- O usuário informa no aplicativo React os ingredientes que possui (ex: "arroz, ovo, tomate").
- O frontend envia uma requisição para a API FastAPI.
- A API utiliza o modelo BERT para encontrar no FAISS as receitas mais relevantes para os ingredientes fornecidos.

- A API enriquece os dados das melhores receitas encontradas e os envia para a API da OpenAI, solicitando que gere uma sugestão amigável e contextual.
- A resposta gerada é retornada pela API ao frontend, que a exibe de forma clara e convidativa para o usuário.

Este fluxo garante uma resposta extremamente rápida na busca e altamente qualitativa e empática na interação, resolvendo o núcleo do problema que seria tornar a dieta algo prazeroso e simples.

Resultados

Apresente resultados do funcionamento do sistema, testes e validações.

4.1 Resultados da Análise de Concorrência

Para fundamentar o desenvolvimento desta solução, foi conduzida uma análise detalhada dos principais aplicativos de nutrição disponíveis no mercado brasileiro, com foco específico em identificar oportunidades estratégicas para atender pessoas com diabetes e usuários com restrições orçamentárias.

4.1.1 Metodologia da Análise

A análise foi realizada através da avaliação sistemática de dados de usuários, incluindo comentários e avaliações em lojas de aplicativos, permitindo uma compreensão profunda das demandas não atendidas pelo mercado atual.

4.1.2 Aplicativos Analisados

- 1. Smart Fit Nutri Resumo: Aplicativo integrado ao ecossistema Smart Fit, focado em avaliação física, bioimpedância e consultas online com nutricionistas Diferenciais: Planos personalizados, acompanhamento profissional e integração com academias Limitações para diabéticos: Não oferece funcionalidades específicas como contagem de carboidratos líquidos, índice glicêmico ou integração com glicosímetros Principais lacunas identificadas: Ausência de lista de compras, substituições nos cardápios, diário alimentar com contagem de macros
- 2. Macros Contador de Calorias Resumo: Focado em contagem de calorias e macronutrientes, com biblioteca de alimentos personalizável Diferenciais: Cálculo de net carbs, biblioteca colaborativa de alimentos Limitações:

Falta de integração com apps de saúde, ausência de modo offline, funcionalidades limitadas para rastreamento de evolução

- 3. YAZIO Resumo: Aplicativo de emagrecimento que combina contador de calorias, jejum intermitente e receitas Diferenciais: Interface moderna, rastreador de jejum, foco motivacional Limitações para diabéticos: Aparentemente sem recursos específicos para diabetes, como índice glicêmico ou monitoramento de glicose
- **4. FatSecret Resumo**: Controle alimentar com foco em comunidade e base extensa de alimentos **Diferenciais**: Versão gratuita robusta, comunidade ativa, integração com Google Fit **Limitações**: Falta de precisão para diabéticos, ausência de net carbs e índice glicêmico, apenas 4 refeições diárias

4.1.3 Principais Oportunidades Identificadas

Lacunas Específicas para Diabéticos: - Ausência de índice glicêmico na maioria dos aplicativos - Falta de integração com monitores de glicose - Ausência de contagem de carboidratos líquidos em vários apps - Controle refinado insuficiente para necessidades metabólicas específicas

Demandas Econômicas Não Atendidas: - Nenhum aplicativo oferece comparação de preços de alimentos - Ausência de funcionalidades de otimização de custo vs. valor nutricional - Falta de sugestões de substituições econômicas mantendo valor nutricional - Inexistência de lista de compras otimizada por custo

Funcionalidades Solicitadas pelos Usuários: - Suporte a dados da TACO (Tabela Brasileira de Composição de Alimentos) - Criação automática de cardápios baseados em preferências e restrições - Lista de compras integrada com planejamento semanal - Reconhecimento de alimentos por IA - Melhor personalização de refeições e horários

4.1.4 Posicionamento Estratégico

A análise revelou uma oportunidade única de mercado: nenhum dos aplicativos analisados combina adequadamente o atendimento às necessidades específicas de pessoas com diabetes com soluções econômicas para otimização de custos alimentares. Esta lacuna representa o diferencial competitivo central do projeto proposto, posicionando-o como pioneiro na intersecção entre saúde especializada e acessibilidade econômica.

4.1.5 Comparação de Aplicativos Concorrentes

4.1.6 1. MyFitnessPal

- Perguntas feitas: peso atual, meta de peso e nível de atividade.
 - Explicações direcionadas para leigos.

- Se você trabalha sentado, é classificado como "pouco ativo", mesmo que treine.
- Sincronizado com o app Health do iPhone.
- Monitora: macros, nutrientes e calorias; adaptável à dieta.
- Seção de receitas com macros detalhados e hashtags como gluten free, dairy free, low carb, low sugar, vegetarian, etc.
- Seção com exercícios para fazer em casa.
- Sincroniza treinos de outros apps (como relógios e apps de celular).
- Mapa com tracking de progresso.
- Possibilidade de criar suas próprias receitas.
- Permite adicionar amigos no app.
- Rastreamento de água.

4.1.7 2. Tecnonutri

- Apresentado por uma assistente virtual.
- Requer aceitar os termos de uso.
- Perguntas feitas: objetivo (emagrecer, ganhar massa), gênero, peso, altura, meta de peso, data de nascimento.
- Dashboard principal mostra contagem de fibras.
- Parceria com o programa Vitat: contato com especialistas gratuitamente.
- Rastreamento de água diária.
- Exercício deve ser adicionado manualmente; não sincroniza com outros apps.
- Possível adicionar foto do prato junto ao cardápio.
- Muitas opções de comida, mas não reconheceu algumas marcas.
- Seção com comunidade para se inspirar em cardápios de outras pessoas.
- Sistema de "XP" por uso contínuo que rende medalhas.
- Acesso a receitas e programas apenas na versão PRO.
- Seção com exercícios para fazer em casa.

4.1.8 3. Programa Mais Saúde (Vitat)

- Apresentado por assistente virtual.
- Requer aceitar os termos de uso.
- Para benefícios, é necessário ter conta no Wellhub (ex-Gympass).
- Perguntas detalhadas sobre:

- Saúde geral (doenças, hábitos, sono, emoções, alimentação, atividade física).
- Ao final, gera uma nota de saúde.
- Parceria com especialistas: análise de refeições via WhatsApp.
- Possível agendar testes e exames pelo app.
- Permite adicionar treinos e refeições.
- Versão PRO:
 - Descontos em farmácias, hospitais e marcas de suplementos.
 - Acesso a comunidades para se inspirar.
 - Agendamento de bioimpedâncias.

4.1.9 4. Alimente-se

- Perguntas feitas: objetivo, nível de atividade, aniversário, gênero, peso, altura, meta de peso, preferências alimentares, número de refeições por dia (mínimo 6), meta em semanas.
- Notificação para beber água.
- Função para comparar macros de dois alimentos diferentes.
- Rastreamento diário de água e exercício.
- Versão PRO:
 - Mostra macros diários.
 - Libera a dieta completa (gratuito mostra só o primeiro dia).
- Área de receitas com filtros por:
 - Dificuldade, tempo, dieta, refeição, macros, ingredientes, utensílios, datas comemorativas, modo de preparo, líquidos.
 - Todas as receitas mostram calorias, mas não os macros.
- Permite adicionar medidas corporais e dobras cutâneas (para % de gordura corporal).

4.1.10 5. Fastic

- Perguntas feitas:
 - Objetivo, gênero, data de nascimento, peso, altura, peso-alvo.

- Nível de atividade, rotina de trabalho, conhecimento sobre nutrição, tentativas anteriores de jejum, doenças, hábitos alimentares, motivações, habilidades na cozinha, frequência de comer fora, etc.
- Requer aceitar os termos de uso.
- Versão PRO:
 - Macros do dia, contador de jejum e receitas.
- Gratuito:
 - Rastreamento de água, passos, peso e exercício.
 - Apenas permite adicionar cardápio diário.
- Interface principal considerada confusa.

4.1.11 6. MyDietCoach

- Requer aceitar os termos de uso.
- Perguntas feitas: objetivo, gênero, altura, peso, peso ideal, nível de atividade, intensidade da dieta, tipo de dieta.
- Permite criar receitas.
- Rastreamento de água.
- Conexão com app Health do iPhone para rastrear exercícios.
- App simples e direto.

4.1.12 7. Diet: WeightLoss

- Perguntas feitas: objetivo, peso-alvo, nível de atividade.
- Sincroniza com o Health do iPhone (importa dados automaticamente).
- Vem com refeições montadas; duas opções por cardápio.
- Não mostra os macros de cada refeição.
- Não permite adicionar refeições próprias.
- Refeições são muito elaboradas, pouco acessíveis à realidade brasileira.
- Quase todas as funções exigem a versão PRO.

4.1.13 8. FitLab

- Conecta automaticamente com o Health do iPhone.
- Perguntas feitas: objetivo (perder gordura, ganhar massa), frequência de treinos, funcionamento intestinal, restrições alimentares, data de nascimento, gênero, peso, altura, peso-alvo.
- Requer cadastro com número de celular.

• Problemas no cadastro: erro ao carregar, não foi possível usar o app.

4.1.14 9. FoodAI

- Perguntas feitas:
 - Através de chatbot: permite perguntas como "Me sinto culpado depois de comer, o que faço?" ou "Tenho metabolismo lento, sugira refeições".
- App fornece dieta já montada com:
 - Ingredientes, modo de preparo, benefícios e valores nutricionais (macros e calorias).
- Substituição de pratos por IA requer versão PRO.
- Registrar cardápio próprio também exige versão PRO.
- Gratuito:
 - Rastreamento de exercício, água e gráficos de evolução.
- Câmera para análise do prato (requer descrição detalhada do conteúdo do prato).

Conclusão

Este projeto emerge como uma solução disruptiva e essencial no cenário nutricional brasileiro, propondo uma abordagem inovadora para combater a dupla carga da má nutrição. Ao invés de focar em restrições e dietas punitivas, a proposta é construir um assistente nutricional baseado em Inteligência Artificial que seja empático, acessível e prático, especialmente para a população de baixa renda. A profunda análise de concorrência revelou uma lacuna crucial no mercado: a ausência de ferramentas que unam a saúde especializada (com foco em condições como o diabetes) à otimização de custos alimentares.

A concepção deste assistente nutricional demonstra que é possível democratizar o acesso à alimentação saudável ao desmistificar mitos e oferecer alternativas realistas. A integração de dados de custo em tempo real com sugestões contextuais e empáticas é a chave para quebrar o ciclo de escolhas alimentares automáticas e impulsionar uma mudança comportamental duradoura. Nós esperamos que a arquitetura modular e escalável, apoiada por tecnologias de ponta como BERT e FAISS, garanta uma experiência fluida e personalizada, transformando o "comer bem" de um privilégio em uma realidade acessível e prazerosa para todos. O sistema tem o potencial de ser um catalisador significativo para a melhoria da saúde pública no Brasil, promovendo bem-estar e dignidade através da alimentação.

5.1 Dificuldades Enfrentadas

Durante o desenvolvimento, algumas dificuldades se destacaram. A coleta e curadoria de dados de receitas brasileiras com a granularidade necessária para o "Funil Inteligente" será um desafio considerável, exigindo robustos scripts de web scraping e heurísticas, para etiquetagem automática, um adendo seria o fato de não conseguir extrair as receitas do TudoGostoso. A criação de um dataset de triplets eficaz para o fine-tuning do BERT, que capture nuances contextuais

e emocionais, exige iterações e validações constantes. Por fim, garantir que as recomendações de custo sejam precisas e atualizadas em tempo real apresenta uma complexidade adicional na integração de fontes de dados dinâmicas.

5.2 Próximos Passos e Sugestões

Para o avanço do projeto, os próximos passos devem focar na validação com usuários reais, já entramos em contato com a coordenadora do atendimento nutricional do CAC para coletar feedbacks e os dados das consultas realizadas (possivelmente teremos esses dados que até o começo do próximo semestre) e precisamos refinar a precisão das recomendações. A integração com plataformas de saúde e mercados locais para otimização de compras pode ser um diferencial estratégico. Por fim, a implementação de funcionalidades específicas para o acompanhamento de condições crônicas, como o diabetes, com o monitoramento de glicose e contagem de carboidratos líquidos, consolidaria ainda mais o valor social e a relevância do assistente.

Referências

As referências estão organizadas no arquivo .
bib e serão exibidas automaticamente.

Kaggle. 2015. "What's Cooking: Recipe Ingredients Dataset." Kaggle competition dataset. https://www.kaggle.com/competitions/whats-cooking.

Ministério da Saúde. 2014. Guia Alimentar Para a População Brasileira. 2nd ed. Brasília, DF: Ministério da Saúde. https://bvsms.saude.gov.br/bvs/publicacoes/guia_alimentar_populacao_brasileira_2ed.pdf.

NEPA - Núcleo de Estudos e Pesquisas em Alimentação, UNICAMP. 2011. "Tabela Brasileira de Composição de Alimentos (TACO)." http://www.nepa.unicamp.br/taco/taco_arquivos/Tabela_TACO_4_edicao_ampliada_e_revisada.pdf.

Universidade CEUB Projeto Integrador I

Emanuela Araújo, Felipe Teodoro, Jamille Rocha, João Vitor, Lucas Wall, Mateo Wall, Victor Vanazzi

Brasília