

Desenvolvimento de um Sistema Inteligente baseado em LLM como Ferramenta de Apoio ao Profissional de Educação Física

João Gabriel Alves Junior¹, Suelen Mapa de Paula¹

¹Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais (IFMG)
Rua Afonso Sardinha, 90 – 36494-018 – Ouro Branco – MG – Brasil

gabrieljunior42566@gmail.com, suelen.mapa@ifmg.edu.br

Abstract. *Balancing individual needs with generic prescriptions is a key challenge in strength training. We present a web-based assistant powered by a fine-tuned GPT-4o-mini model that produces evidence-based workout plans for physical-education professionals. Trained on ACSM and NSCA guidelines and deployed with a React/Node/MySQL stack, the system was reviewed by certified coaches, who judged the plans high-quality and requiring only minor edits. These results suggest that a domain-adapted LLM can streamline prescription work while preserving professional judgment and client interaction.*

Keywords: *Generative Artificial Intelligence; Strength Training; Intelligent System; Chat Assistant; Training Personalization.*

Resumo. *A personalização do treinamento de força ainda enfrenta limitações decorrentes de prescrições genéricas, o que pode ser um dos fatores para a baixa adesão da população à prática regular de atividade física. Este trabalho apresenta um assistente web, apoiado em um modelo GPT-4o-mini, ajustado com diretrizes do ACSM e da NSCA, que gera planos de treino baseados em evidências para profissionais de Educação Física. Construído com React, Node.js e MySQL, o sistema foi avaliado por especialistas, que classificaram os programas gerados como seguros e relevantes, exigindo apenas ajustes pontuais. Os resultados indicam que um LLM especializado pode aumentar a eficiência da prescrição, sem substituir o julgamento crítico nem a interação humana.*

Palavras-chave: *Inteligência Artificial Generativa, Treinamento de Força, Sistema inteligente, Chat Assistente, Personalização de Treino.*

1. Introdução

Programas de treinamento de força deixaram de ser privilégio de atletas de elite para se consolidarem como pilar de saúde pública. Em 2023, o mercado global de academias movimentou US\$ 98,1 bilhões e pode atingir US\$172,9 bilhões até 2028 [mor 2023]. No Brasil, o *Panorama Setorial Fitness Brasil 2024* registrou 25602 academias cadastradas, mas apenas 5% da população as frequenta regularmente [fit 2024]. O contraste entre a oferta crescente e a baixa participação reflete ainda a inobservância das recomendações da Organização Mundial da Saúde: 91,4% dos adultos brasileiros não cumprem os dois dias semanais de exercícios de resistência indicados [Stopa et al. 2023].

Além da baixa adesão, a retenção inicial é crítica. Em um grupo de 5240 novos matriculados em uma rede de academias no Rio de Janeiro, 63% abandonaram o programa antes de 90 dias e 96% não permaneceram após 12 meses [Sperandei et al. 2016].

Estudos atribuem grande parte desse abandono às prescrições genéricas, a chamada “ficha padrão”, que ignoram histórico clínico, preferências e objetivos individuais, aumentando o risco de lesões e diminuindo a motivação [Carben et al. 2022; Liz et al. 2010].

Diante desse cenário, o presente trabalho desenvolveu e avaliou um sistema inteligente baseado em um Modelo de Linguagem de Grande Escala (em inglês, *Large Language Model* – LLM) especializado, capaz de auxiliar profissionais de Educação Física na elaboração de planos de treino personalizados, e alinhados às diretrizes de referência do ACSM (*American College of Sports*¹ – Colégio Americano de Medicina Esportiva) e da NSCA (*National Strength and Conditioning Association*² – Associação Nacional de Força e Condicionamento).

A proposta buscou tratar dois problemas específicos deste cenário:

- o tempo limitado dos profissionais frente a grandes carteiras de alunos;
- a necessidade de planos de treinos individualizados, fundamentados em evidências científicas.

Para apresentar de forma clara o desenvolvimento e a análise do sistema proposto, este artigo está estruturado da seguinte forma. A Seção 2 revisa os trabalhos correlatos na área de aplicação de tecnologia ao treinamento. A Seção 3 detalha a metodologia, incluindo a base de conhecimento científico utilizada, a arquitetura técnica do sistema e o protocolo de validação. A Seção 4 apresenta e discute os resultados obtidos, analisando a eficácia do sistema. Finalmente, a Seção 5 conclui o trabalho, sintetizando as contribuições e delineando direções para pesquisas futuras.

2. Revisão da Literatura

Para contextualizar o sistema desenvolvido neste trabalho, esta seção analisa publicações que investigaram a aplicação de sistemas computacionais na personalização do treinamento físico. Foram selecionados trabalhos que ilustram a evolução das abordagens tecnológicas ao longo do tempo, desde os sistemas especialistas clássicos, passando por sistemas de recomendação com aprendizado de máquina, até as mais recentes aplicações de Modelos de Linguagem de Grande Escala (LLMs). A análise busca destacar os diferenciais e a lacuna que o presente trabalho visa preencher, dando ênfase em 5 artigos que foram considerados os mais relevantes.

2.1. Trabalhos Correlatos

A primeira onda de automação na prescrição de treinos se apoiou em sistemas especialistas tradicionais, que operam com base em um conjunto de regras “se-então”. Um exemplo representativo foi o sistema Ex-Pres, desenvolvido por [Bastos et al. 2007] para prescrever exercícios a indivíduos com hipertensão. Essa metodologia, que consistia em codificar o conhecimento de especialistas em uma base de regras fixas, demonstrou ser funcional para nichos bem definidos. Contudo, sua principal limitação reside na rigidez: tais sistemas são incapazes de lidar com perfis de usuários complexos ou com objetivos não previstos em suas regras. Adicionalmente, a manutenção e a escalabilidade se mostraram problemáticas, pois qualquer nova exceção ou conhecimento exigia programação manual.

¹[of Sports Medicine 2022]

²[Strength and Association 2021]

Com a evolução da Inteligência Artificial, a abordagem seguinte foi marcada pelos sistemas de recomendação que utilizam algoritmos de *Machine Learning*. O trabalho de [Valcarce-Espinosa et al. 2021], por exemplo, propôs um sistema para recomendar atividades físicas com base no perfil do usuário e no comportamento de perfis similares, usando técnicas como a filtragem colaborativa. Embora mais dinâmicos que os sistemas baseados em regras, eles apresentam falhas críticas no contexto da prescrição de treinamento de força. Primeiramente, o problema de “partida a frio” dificulta a geração de recomendações para novos usuários. De forma mais crucial, suas sugestões são baseadas em correlações estatísticas (o que pessoas parecidas fazem) e não em um entendimento profundo de princípios fisiológicos e biomecânicos específicos de um indivíduo. Isso pode levar a recomendações populares, mas que são subótimas ou potencialmente inseguras para o indivíduo.

A mais recente e promissora onda de inovação vem da aplicação de LLMs. Pesquisadores têm explorado esses modelos para criar sistemas de *coaching* mais sofisticados. Um exemplo notável é o “GPTCoach”, de [Matthew Jörke 2025], que buscou superar as limitações de um LLM genérico ao integrá-lo com um programa de saúde baseado em evidências e estratégias de aconselhamento motivacional. Este tipo de trabalho evidencia um movimento importante para especializar os LLMs. Abordagens que utilizam modelos de linguagem de forma genérica, sem um processo de curadoria ou especialização, correm o risco significativo de gerar informações desatualizadas ou desalinhadas com o consenso científico, dada a natureza de seu treinamento em dados vastos e não específicos ao domínio do exercício.

Pesquisas mais recentes reforçam essa tendência. [Lim et al. 2025] apresentaram o *Micha*, um chatbot baseado no GPT-4 que incorpora princípios de Entrevista Motivacional para apoiar mudanças de comportamento em saúde. Em testes com 34 participantes, o agente obteve pontuações de usabilidade (SUS=83,1) superiores ao corte de excelência e aderiu a 92% das boas-práticas de aconselhamento, mas os autores ressaltam a necessidade de maior personalização fisiológica para contextos específicos, como o treinamento de força.

Complementarmente, [Ong et al. 2024] compararam respostas de um LLM genérico e de *health coaches* humanos a 60 perguntas sobre sono e bem-estar. O modelo igualou ou superou os especialistas em empatia e clareza. Porém, mostrou lacunas quando solicitado a propor intervenções concretas baseadas em evidências. O estudo conclui que, embora promissores, os LLMs requerem domínio adaptado (um termo comum da área e conhecido como: *domain-tuning*) para recomendações seguras em cenários clínicos ou de performance física.

Diante do exposto, o sistema desenvolvido neste trabalho se posiciona de forma diferenciada. Como detalhado no quadro 1, ele avança em relação às abordagens anteriores ao combinar a flexibilidade semântica de um LLM com a robustez do conhecimento científico explícito. Diferentemente dos sistemas rígidos baseados em regras [Bastos et al. 2007], interpreta perfis complexos; em contraste com os sistemas de recomendação [Valcarce-Espinosa et al. 2021], fundamenta suas prescrições em princípios fisiológicos, e não em correlações estatísticas. Na comparação com aplicações recentes de LLMs focadas em mudança comportamental, como o *GPTCoach* de [Matthew Jörke 2025], o chatbot *Micha* com Entrevista Motivacional [Lim et al.

2025] e o modelo genérico avaliado por [Ong et al. 2024], o diferencial central do sistema proposto reside na metodologia de ajuste fino (do inglês *fine-tuning*) sobre um conjunto de dados, calibrado a partir das diretrizes do ACSM e da NSCA. Esse processo eleva a confiabilidade e o alinhamento das recomendações com o padrão-ouro da área, suprimindo as lacunas de personalização fisiológica apontadas nesses estudos. Por fim, a solução foi concebida desde o início para atuar como um ajudante digital. Isto é, uma ferramenta de apoio que potencializa a capacidade e a eficiência do especialista humano, preenchendo a lacuna de suporte profissional identificada nos trabalhos analisados.

Quadro 1: Análise comparativa das abordagens e o posicionamento do sistema proposto.

Trabalho	Abordagem	Principais Limitações	Diferencial do Sistema Proposto
[Bastos et al. 2007]	Sistema Especialista (Baseado em Regras)	Rigidez, baixa escalabilidade, dificuldade em lidar com perfis complexos.	Flexibilidade semântica para interpretar dados complexos do usuário.
[Valcarce-Espinosa et al. 2021]	Machine Learning (Filtragem Colaborativa)	Baseado em correlação (popularidade), não em fisiologia; problema de <i>cold-start</i> ; risco de sugestões inseguras.	Prescrição fundamentada em conhecimento científico explícito, garantindo segurança e eficácia.
[Matthew Jörke 2025]	LLM com integração de conhecimento externo	Risco de desalinhamento se não for especializado; foco no usuário final, não no profissional.	<i>Fine-tuning</i> com <i>dataset</i> curado de diretrizes científicas para alta confiabilidade; foco no apoio ao profissional.
[Lim et al. 2025]	LLM + Entrevista Motivacional	Foco em mudança comportamental genérica; falta de parâmetros fisiológicos.	Mostra viabilidade de MI em LLM, mas carece de prescrição de força baseada em diretrizes.
[Ong et al. 2024]	LLM genérico vs. health coaches	LLM empático porém impreciso em intervenções específicas.	Evidencia necessidade de <i>domain-tuning</i> para qualidade profissional.

Continuação do quadro 1 da página anterior

Trabalho	Abordagem	Principais Limitações	Diferencial do Sistema Proposto
Sistema Desenvolvido	LLM com <i>Fine-Tuning</i>	Dependência da qualidade da informação de entrada (GIGO); natureza “caixa-preta” com explicabilidade limitada; validação restrita a amostra inicial.	Combina flexibilidade semântica, conhecimento explícito e é projetado como um ajudante para o especialista humano.

3. Metodologia

Esta seção detalha as três fases cruciais do sistema proposto:

1. Base de Conhecimento: a construção da base de conhecimento científico que fundamenta a inteligência artificial;
2. Arquitetura do Sistema: apresenta a implementação da arquitetura técnica que transforma o conhecimento em uma solução funcional;
3. Protocolo de Validação: trata do delineamento de um protocolo multifacetado para aferir a eficácia e a segurança do sistema.

Elas serão detalhadas na seções que se seguem.

3.1. Base de Conhecimento

A eficácia e a segurança de qualquer sistema inteligente de prescrição dependem da qualidade e profundidade do conhecimento que o alimenta. Portanto, a primeira etapa metodológica consistiu em estabelecer dados científicos (“verdade fundamental”) que servem como alicerce para a inteligência do sistema. Este conhecimento não foi apenas baseado em diretrizes gerais, mas sim extraído e traduzido a partir de uma revisão sistemática de meta-análises e estudos seminais da literatura científica, que definem o estado da arte na ciência do exercício.

A partir deste estudo aprofundado, a base de conhecimento foi estruturada em pilares interdependentes que se tornaram os parâmetros e a lógica operacional do modelo de IA. Estes pilares incluem: a avaliação individual, os princípios cardeais do treinamento (como sobrecarga progressiva e especificidade e, mais importante, a manipulação das variáveis de treino). O Quadro 2 detalha como a evidência científica dos trabalhos foi operacionalizada para calibrar o comportamento do modelo, garantindo que cada prescrição seja segura, eficaz e altamente personalizada.

Quadro 2: Detalhamento da Base de Conhecimento para Calibração do Modelo de IA.

Fonte / Conceito	Conhecimento Científico Extraído	Função e Implementação no Modelo
Tensão Mecânica [Schoenfeld 2010]	A tensão mecânica é o principal motor da hipertrofia. Estresse metabólico e dano muscular são fatores secundários.	Lógica Central: O modelo prioriza exercícios e esquemas de carga que maximizem tensão mecânica em amplitude completa.
Volume de Treino [Baz-Valle et al. 2022]	Faixa de 12–20 séries semanais por grupo muscular otimiza hipertrofia, com variações específicas por músculo.	Cálculo de Volume: Usa essa faixa para definir o volume total e distribui conforme frequência semanal.
Frequência de Treino [Schoenfeld et al. 2019]	Com volume igualado, frequência (1–3+ vezes/semana) não altera muito a hipertrofia.	Ferramenta Logística: Organiza o volume em treinos full-body, ABC, ABCDE etc., conforme agenda do usuário.
Seleção de Cargas [Schoenfeld et al. 2017]	Altas (baixas reps) vs. baixas (altas reps) geram hipertrofia similar próximo à falha; altas são melhores para força.	Flexibilidade de Prescrição: Recomenda zonas de repetições (ex.: 6–10 ou 20–25) adaptadas a preferências ou limitações.
Proximidade da Falha [Grgic et al. 2022a]	Treinar a 1–3 repetições na reserva é tão eficaz quanto falhar, gerando menos fadiga.	Gerenciamento de Fadiga: Por padrão, usa RIR; falha só para perfis avançados ou exercícios de isolamento.
Intervalos de Descanso [Singer et al. 2024]	Descansos maiores que 90s favorecem hipertrofia ao manter maior volume total.	Otimização de Performance: Prescreve 120–180 s em compostos, justificando com manutenção de performance.
Seleção de Exercícios [Grgic et al. 2022b]	Multiarticulares e monoarticulares são equivalentes em hipertrofia com volume equalizado; multiarticulares são mais eficientes.	Estratégia de Exercícios: Base no multiarticular eficiente, com isolations adicionais para simetria e cobertura completa.

3.2. Arquitetura do Sistema

A transformação da base de conhecimento científico em um sistema funcional foi realizada através do desenvolvimento de uma aplicação *web full-stack*. Essa abordagem encapsula a inteligência do modelo de IA em uma interface acessível e intuitiva para o profissional de Educação Física, conforme ilustra sua tela inicial na Figura 1.

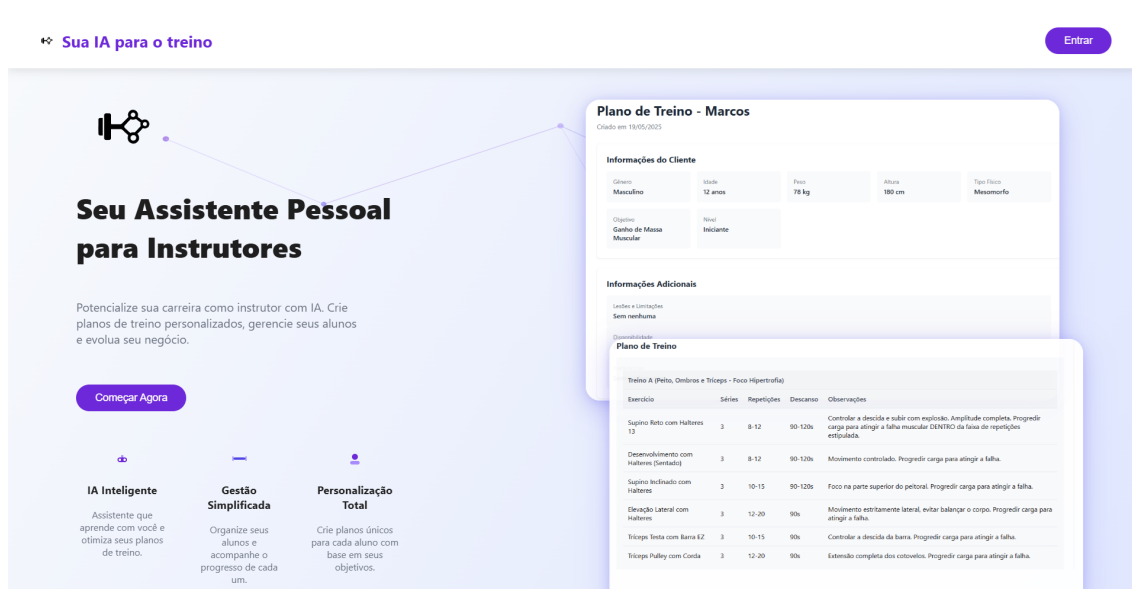


Figura 1. Landing page de apresentação do sistema.

A sistema está disponível para demonstração, mediante registro prévio a partir do seguinte link de acesso: link de acesso ao sistema³.

3.2.1. Processo de *Fine-Tuning* e Geração de Dados

O processo de *fine-tuning* transforma um modelo de linguagem de propósito geral, como o GPT-4o-mini, em um especialista focado em uma tarefa específica, neste caso, a prescrição de treinamento de força. O objetivo desta etapa foi ensinar ao modelo tanto o formato de saída desejado (tabelas em HTML) quanto, e mais importante, a aderir estritamente aos princípios científicos e diretrizes de segurança definidos no Quadro 2.

Para isso, foi criado um conjunto de dados (*dataset*) de exemplos de alta qualidade em formato *JSON Lines* (JSONL). Cada linha representa um diálogo completo com três papéis: *system*, que define a persona da IA como um “assistente especialista”; *user*, que traz o perfil detalhado do cliente; e *assistant*, que contém a resposta ideal. Tanto os cenários do usuário quanto as respostas foram gerados artificialmente utilizando o *Gemini 2.5 Pro*. Tanto para criar perfis fictícios variados e depois, em uma instância separada e configurada com as mesmas diretrizes científicas, para produzir as prescrições correspondentes⁴. Essa estratégia permitiu controlar a qualidade sem demandar rotula-

³<https://gym-ai-one.vercel.app/>

⁴Detalhes sobre a geração automática dos diálogos podem ser consultados em [Google DeepMind 2025].

gem manual extensiva, além de garantir aderência às evidências, como priorizar volume para hipertrofia [Schoenfeld et al. 2017] ou repetições baixas para força máxima [Haff and Triplett 2016]. A Figura 3.2.1 ilustra um exemplo de entrada no *dataset*.

O ajuste foi realizado via API da OpenAI, calibrando hiperparâmetros – como número de épocas – para otimizar o aprendizado e evitar *overfitting*, resultando em um modelo customizado e especializado.

```
{ "messages": [
  { "role": "system", "content": "Você é um assistente
  especialista..." },
  { "role": "user", "content": "Nome: Gabriel Lima, Gênero:
  Masculino,
  Idade: 24, ..., Objetivo: Hipertrofia..., Nível:
  Intermediário..." },
  { "role": "assistant", "content": "<table>...Treino A...
  </table>..." }
]}
```

Figura 2. Exemplo de uma instância no conjunto de dados de fine-tuning (formato JSONL)⁶

3.2.2. Arquitetura Full-Stack e Funcionalidades

O sistema foi implementado com uma arquitetura modular para garantir escalabilidade e manutenibilidade, composta pelos seguintes elementos:

- **Frontend:** Desenvolvido com *React.js*, é responsável por toda a interface com o usuário, incluindo formulários para entrada de dados, a exibição dos planos de treino (Figura 3) e a interface do *chat* interativo (Figura 4).
- **Backend:** Construído com *Node.js* e o ORM Prisma, atua como o cérebro da aplicação. Ele orquestra o fluxo de dados, constrói os *prompts*, comunica-se com a API da OpenAI e persiste as informações.
- **Banco de Dados:** Utilizando *MySQL*, armazena de forma segura os perfis dos clientes, históricos e todos os planos de treino gerados para consulta futura.
- **Inteligência Artificial:** A API da OpenAI é o componente que processa os *prompts* através do modelo previamente *fine-tuned* para gerar as prescrições.

⁶O conteúdo foi abreviado para fins de visualização.

Plano de Treino - Ficha 3

Criado em 06/06/2025

Informações do Cliente

Gênero Masculino	Idade 60 anos	Peso 70 kg	Altura 170 cm	Tipo Físico Ectomorfo
Objetivo Resistência	Nível Iniciante			

Informações Adicionais

Lesões e Limitações
Ossos fracos, maior idade

Disponibilidade
3x na semana

Preferências
Treinos curtos

Plano de Treino

Treino A (Corpo Inteiro - Foco Endurance)				
Exercício	Séries	Repetições	Descanso	Observações

Figura 3. Exibição de um plano de treino gerado.

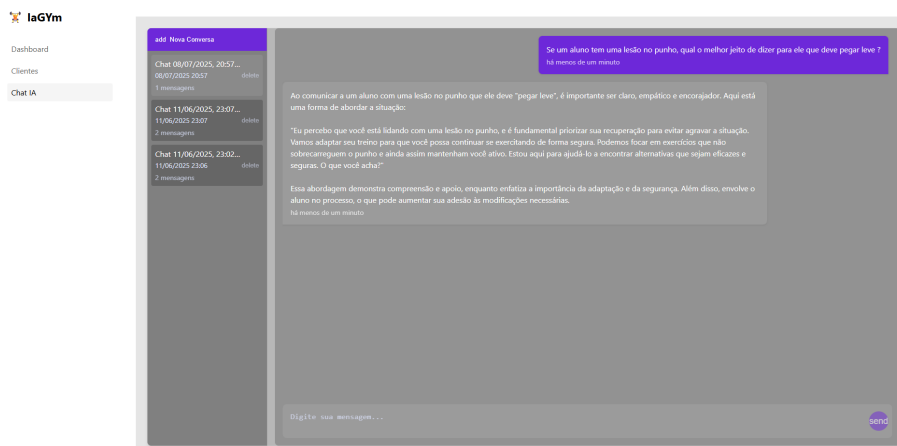


Figura 4. Interface do chat para consulta interativa.

O fluxo de dados do sistema, ilustrado na Figura 5, inicia-se no *frontend*, onde o profissional insere as informações do cliente. O *backend* recebe estes dados, formata-os em um *prompt* estruturado e os envia para a API da OpenAI. A resposta, contendo o plano de treino, é recebida, armazenada no banco de dados *MySQL* e, finalmente, enviada de volta ao *frontend* para ser exibida ao profissional.

3.3. Protocolo de Validação

Para aferir a robustez, eficácia e o alinhamento do sistema, foi delineado um protocolo de validação qualitativa, centrado na avaliação por especialistas de domínio. Foram recrutados dois profissionais de Educação Física certificados, com experiência em prescrição de treinamento de força, para atuarem como avaliadores. A eles, foi solicitado que interagis-

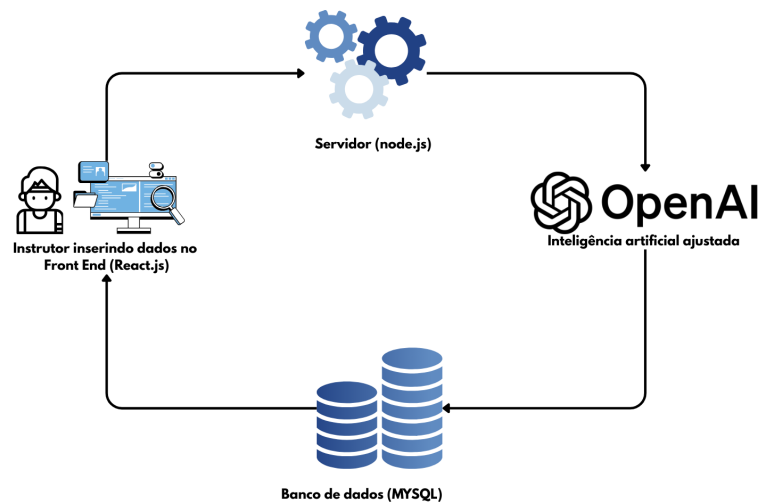


Figura 5. Ilustração do fluxo de dados da aplicação, desde a entrada do profissional até a exibição do plano de treino gerado pela IA.

sem com o sistema e avaliassem seus resultados em três cenários distintos, para testar o desempenho sob diferentes condições:

1. **Cenário 1:** Simulação de Operador com Conhecimento Limitado. Os especialistas inseriram *prompts* de usuário genéricos (ex: objetivo “hipertrofia”, sem detalhes sobre lesões). O objetivo era verificar se o sistema geraria um plano de treino seguro e coerente, mesmo com informações de entrada incompletas, simulando um operador menos experiente.
2. **Cenário 2:** Simulação de Operador Especialista. Em seguida, os especialistas forneceram *prompts* ricos em detalhes e contexto, incluindo objetivos específicos, histórico de lesões e preferências, replicando uma avaliação completa. O objetivo era testar a capacidade do sistema de gerar um plano altamente personalizado e otimizado.

Para cada plano de treino gerado nos três cenários, os especialistas forneceram uma avaliação qualitativa detalhada, focando nos critérios de Segurança (ausência de exercícios contraindicados), Eficácia Científica (alinhamento com os princípios do treinamento e evidências recentes) e Aplicabilidade Prática (viabilidade da execução em um ambiente de academia comercial).

4. Resultados e Discussões

Esta seção apresenta os resultados obtidos a partir do protocolo de validação por especialistas e discute suas implicações, conectando os achados com a tese central do trabalho: o sistema de IA como uma ferramenta de aumento de capacidade para o profissional de Educação Física.

Embora a validação tenha sido realizada com um número limitado de profissionais, não possuindo representatividade estatística para generalizações amplas, os resultados apurados foram promissores e corroboraram a hipótese do projeto, revelando uma performance distinta do sistema em cada um dos cenários propostos.

No cenário com prompts genéricos, os especialistas observaram que os planos de treino gerados pelo sistema mantiveram um alto padrão de segurança e coerência com as diretrizes gerais. Os planos foram classificados como “bons” e “seguros”. Contudo, o avaliador destacou que os treinos, embora corretos, “fogem de alguns detalhes que o treinador em si sabe”, como a inclusão de exercícios corretivos específicos ou variações de cadência. Em suma, o sistema atuou como uma “grade de proteção”, fornecendo um sistema básico, porém seguro e de qualidade, mas sem o refinamento que apenas a experiência humana pode prover.

Quando os especialistas forneceram dados de entrada detalhados, a qualidade dos planos gerados foi classificada como “excelente”. Os avaliadores reportaram que seu trabalho subsequente se resumiu a “apenas refinar os detalhes”, como a alteração na ordem de um exercício acessório. Concluiu-se que o sistema automatizou com sucesso entre 90% a 95% do trabalho de estruturação do plano, liberando o profissional para a etapa final de personalização de alto nível.

4.1. Discussão dos Resultados e Validação da Tese

Os resultados validam a proposta deste trabalho: o sistema de IA atua como um verdadeiro ajudante, e não como um substituto autônomo. A diferença de performance entre os cenários ilustra o princípio fundamental “*Quality In, Quality Out*”, reforçando o papel central e insubstituível do profissional no processo.

Para o profissional menos experiente ou sobrecarregado (Cenário 1), o sistema serve como um apoio de segurança, mas com qualidade. Ele garante que a base da prescrição seja cientificamente sólida, prevenindo erros graves e estabelecendo um padrão de excelência.

Para o profissional experiente (Cenário 2), a ferramenta se transforma em um catalisador de produtividade e eficiência. Ao automatizar a tarefa demorada de estruturar o plano, libera o tempo do especialista para focar onde é mais valioso. Isto é, na interpretação de nuances, na aplicação da experiência prática e na interação humana e motivacional.

Portanto, a validação demonstra que o sistema atinge seu objetivo principal de formalizar e escalar o conhecimento especializado, permitindo que os profissionais entreguem um serviço de maior qualidade e de forma mais eficiente, solidificando o modelo de colaboração Homem e IA.

5. Conclusão

Este trabalho apresentou o desenvolvimento, a arquitetura e a validação de um sistema inteligente projetado para atuar como um assistente especialista para profissionais de Educação Física. Ao integrar conhecimento científico em um Modelo de Linguagem de Grande Escala (LLM) especializado via *fine-tuning*, a solução buscou endereçar as limitações de eficiência e padronização dos métodos de prescrição manuais.

Com isso, as contribuições deste estudo podem ser sintetizadas em três áreas fundamentais.

- Técnica: Foi demonstrada a viabilidade da construção de um sistema inteligente funcional que utiliza um LLM *fine-tuned* para gerar planos de treino comple-

xos, incluindo uma interface de *chat* interativa. A arquitetura *full-stack* (React, Node.js, MySQL) provou ser uma base robusta e escalável.

- Validação Empírica da ideia “IA como ajudante”: Os resultados confirmaram que o sistema aumenta a capacidade profissional em vez de substituí-la. Para o operador com informações limitadas, atuou como um “apoio de segurança”. Para o operador especialista, funcionou como um “catalisador de produtividade”, automatizando o trabalho de base e liberando o profissional para o refinamento de alto nível.
- Metodológica: O projeto oferece um *framework* para o desenvolvimento de ferramentas de IA em domínios da saúde. A abordagem de formar um *dataset* baseado em evidências, realizar o *fine-tuning* e validar com especialistas em cenários de uso real pode servir como um roteiro para futuros trabalhos.

Entretanto, reconhece-se que o sistema possui limitações. A principal é a dependência da qualidade da informação de entrada (*Quality In, Quality Out*). A natureza “caixa-preta” dos LLMs impõe desafios à explicabilidade, e a validação em pequena escala aqui apresentada e necessita de estudos mais abrangentes para generalização dos resultados. Além disso, se faz necessário ampliar o cenário de teste, ouvindo o *feedback* de outros profissionais. Essas limitações, contudo, abrem caminhos promissores para trabalhos futuros:

- Validação em Larga Escala: Conduzir um estudo com um grupo maior de profissionais e clientes para medir o impacto em indicadores de performance, satisfação e aderência ao treinamento a longo prazo.
- Expansão do *Dataset*: Enriquecer o conjunto de dados com perfis de populações diversas (idosos, gestantes, com comorbidades) para mitigar vieses e ampliar a aplicabilidade.
- Implementação de Ciclo Fechado (*Closed-Loop*): Permitir que o praticante insira *feedback* (ex: Percepção Subjetiva de Esforço - RPE) para que o sistema ajuste dinamicamente os planos futuros, individualizando a sobrecarga progressiva.
- Aprimoramento da Explicabilidade: Focar em técnicas que permitam ao sistema justificar suas escolhas, fortalecendo seu papel como ferramenta de aprendizado.

Sendo assim, conclui-se que este trabalho reforça o potencial da inteligência artificial como uma aliada valiosa e responsável na promoção da saúde, evidenciando um caminho promissor para ampliar a expertise humana e aprimorar a qualidade da prescrição de exercícios.

Referências

- (2023). Health and fitness club market — growth, trends, and forecasts (2024–2028). Acesso em julho 2025.
- (2024). Panorama setorial fitness brasil 2024. Acesso em julho 2025.
- Bastos, W. G., Pimenta, M. B., and Pithon, A. J. (2007). Ex-pres: an expert system for exercise prescription for hypertensive individuals. In *2007 IEEE International Symposium on Industrial Electronics*, pages 2693–2698. IEEE.
- Baz-Valle, E., Balsalobre-Fernández, C., Alix-Fages, C., and Santos-Concejero, J. (2022). A systematic review of the effects of different resistance training volumes on muscle hypertrophy. *Journal of Human Kinetics*, 81(1):199–210.

- Carben, J. T., Eannucci, J. F., Zourdos, M. C., Bittenbender, T. M., Ormsbee, M. J., Gabor, K. A., and White, J. A. (2022). The efficacy of a 'one-size-fits-all' exercise program on anthropometric and physical fitness variables among law enforcement officers. *Work*, 72(1):113–122.
- Google DeepMind (2025). Gemini 2.5 pro technical overview. <https://deepmind.google/models/gemini/pro/>; acesso em julho 2025.
- Grgic, J., Schoenfeld, B. J., et al. (2022a). Effects of resistance training performed to repetition failure or non-failure on muscular strength and hypertrophy. *Journal of Sport and Health Science*, 11(2):175–184.
- Grgic, J., Schoenfeld, B. J., et al. (2022b). Hypertrophic effects of single-versus multi-joint exercise of the limb muscles: A systematic review and meta-analysis. *Sports Medicine*, 52(3):683–700.
- Haff, G. G. and Triplett, N. T. (2016). *Essentials of Strength Training and Conditioning*. Human Kinetics, 4th edition.
- Lim, S.-M., Lee, K., and Chen, V. (2025). Llm-based conversational agents for behaviour change support: The “micha” chatbot. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*. Avalia um chatbot GPT-4 com Entrevista Motivacional para promover atividade física.
- Liz, C. M., Crocetta, T. B., Viana, M. d. S., Brandt, R., and Andrade, A. (2010). Aderência à prática de exercícios físicos em academias de ginástica. *Motriz. Revista de Educação Física*, 16(1):181–188.
- Matthew Jörke, Shardul Sapkota, L. W. N. V. P. S. E. B. J. A. L. (2025). GPTCoach: Towards LLM-Based Physical Activity Coaching.
- of Sports Medicine, A. C. (2022). *ACSM's Guidelines for Exercise Testing and Prescription*. Wolters Kluwer, 11 edition.
- Ong, Q. C., Ang, C.-S., and Chee, D. Z. Y. e. a. (2024). Advancing health coaching: a comparative study of large language model and health coaches. *Artificial Intelligence in Medicine*, 157:103004.
- Schoenfeld, B. J. (2010). The mechanisms of muscle hypertrophy and their application to resistance training. *Journal of Strength and Conditioning Research*, 24(10):2857–2872.
- Schoenfeld, B. J., Grgic, J., and Krieger, J. (2019). How many times per week should a muscle be trained to maximize muscle hypertrophy? *Journal of Sports Sciences*, 37(11):1286–1291.
- Schoenfeld, B. J., Grgic, J., Ogborn, D., and Krieger, J. W. (2017). Strength and hypertrophy adaptations between low- vs. high-load resistance training: a systematic review and meta-analysis. *The Journal of Strength & Conditioning Research*, 31(12):3508–3523.
- Singer, A., Wolf, M., Generoso, L., and ... Schoenfeld, B. J. (2024). Give it a rest: a systematic review with Bayesian meta-analysis on the effect of inter-set rest interval duration on muscle hypertrophy. *Frontiers in Sports and Active Living*, 6:1429789.
- Sperandei, S., Vieira, M. C., and Reis, A. C. (2016). Adherence to physical activity in an unsupervised setting: explanatory variables for high attrition rates among fitness center members. *Journal of Science and Medicine in Sport*, 19(11):916–920.
- Stopa, S. R., Szwarcwald, C. L., Oliveira, M. M., Gouvêa, E. C. D. P., Vieira, M. L. F. P., Freitas, M. P. S., and et al. (2023). Perfil de atividade física no tempo livre e comportamento sedentário em adultos no brasil: pesquisa nacional, 2019. *Epidemiologia e Serviços de Saúde*, 32(2):e2021309.

Strength, N. and Association, C. (2021). *NSCA's Essentials of Personal Training*. Human Kinetics, 3 edition.

Valcarce-Espinosa, D., Llauradó, O., and Pardàs, F. (2021). A hybrid recommender system for physical activity and health. *IEEE Access*, 9:116900–116918.