



Conecte-se com o CI-IA

<https://linktr.ee/iasaudeufmg>



Avaliação aponta impacto positivo do “Curso de Proteção de Dados em Pesquisa em Saúde”

O “Curso de Proteção de Dados em Pesquisa em Saúde”, promovido pelo CI-IA Saúde entre **2 e 6 de junho de 2025**, recebeu **avaliação altamente positiva**. Realizado online, o curso teve 50 inscritos, dos quais 34 responderam à pesquisa de avaliação, revelando resultados expressivos quanto à relevância e aplicabilidade da formação.

Segundo os dados coletados, 88% dos participantes avaliaram o conteúdo como “Relevante” ou “Muito Relevante” para seu aprimoramento profissional. Além disso, 94% afirmaram que conseguem aplicar claramente o conhecimento adquirido

em atividades que já desenvolvem no dia a dia, evidenciando o alinhamento do curso com as demandas reais de profissionais e estudantes da saúde e áreas afins.

As respostas abertas reforçam a percepção positiva. Participantes relataram que, mesmo fora do campo da pesquisa, os temas tratados — como Lei Geral de Proteção de Dados (LGPD), Autoridade Nacional de Proteção de Dados (ANPD) e estudos de caso — trouxeram maior clareza e preparo para lidar com dados sensíveis. Exemplos de aplicação incluem a elaboração de documentos com anonimização de dados, propostas de pesquisa em grupos acadêmicos, práticas de segurança em organizações e a condução ética de processos clínicos e administrativos.

Além da qualidade do conteúdo, o curso também se destacou pelo engajamento dos participantes: **97% das avaliações já realizadas indicam que o nível de participação geral foi considerado Satisfeito ou Muito Satisfeito**.

A avaliação comprova o êxito da iniciativa do CI-IA Saúde em promover formação qualificada e ética sobre o tratamento de dados em saúde, reafirmando o compromisso com as capacitações.

Conheça os projetos vigentes no CI-IA

“O uso da Inteligência Artificial para predição de sobrepeso e obesidade em crianças brasileiras”

Universidade Federal de Minas Gerais

Jorge Gustavo Velasquez
Melendez
Coordenador

Objetivo do Projeto

Este trabalho tem como objetivo o desenvolvimento de um modelo preditivo de classificação, para antecipar casos de obesidade infantil em crianças de 0 a 9 anos, utilizando como ponto de referência os dados disponíveis até os 5 anos de idade. A previsão da obesidade foi baseada em uma base de dados abrangente, que incluiu informações maternas, características da gestação, dados paternos, variáveis temporais relacionadas ao crescimento das crianças, além de medidas como peso, altura e histórico de doenças.

Metas

- Estabelecer fatores que afetam as trajetórias do IMC (Índice de Massa Corporal) em crianças de 0 a 10 anos.
- Identificar como condições do binômio mãe-filho e do período pré-natal predizem o pico e o menor valor de IMC

ao longo da infância.

– Avaliar o valor preditivo da obesidade aos 9 anos com base em variáveis desde o nascimento.

A metodologia adotada seguiu uma abordagem sistemática de ciência de dados, iniciando com a preparação e análise exploratória dos dados. Foram realizadas análises estatísticas e gráficas para entender o comportamento das variáveis e sua relação com a obesidade. A partir disso, foram construídos e testados modelos de *machine learning* com foco na performance preditiva.

Resultados parciais

Com base em uma coorte de cerca de seis milhões de crianças brasileiras de baixa renda, a pesquisa utiliza técnicas de aprendizado de máquina para análise preditiva em larga escala. Já foi produzido um manuscrito científico pronto para submissão, com achados iniciais sobre os determinantes precoces da obesidade. Um dos desafios é o processo de *data linkage*, que demanda infraestrutura computacional e equipe qualificada.

Algoritmos Testados

Três algoritmos principais foram aplicados e comparados:

Decision Tree: Utilizado como modelo base pela sua interpretabilidade, permitindo identificar as variáveis mais relevantes para a obesidade infantil.

Random Forest: Um modelo de conjunto que melhora a robustez e a generalização ao reduzir o risco de



PARCERIA



FINANCIADORES



COORDENAÇÃO



overfitting, agregando múltiplas árvores de decisão.

XGBoost: Algoritmo de *gradient boosting* altamente eficiente, conhecido por sua capacidade de lidar com dados desequilibrados e capturar interações complexas entre variáveis.

Avaliação dos Modelos

A principal métrica de avaliação utilizada foi a curva ROC (*Receiver Operating Characteristic*), que mede a capacidade dos modelos de distinguir entre classes (obeso vs. não obeso) de forma balanceada, considerando sensibilidade e especificidade. A análise da Área Sob a Curva ROC (AUC) permitiu comparar o desempenho entre os modelos de forma precisa e padronizada.

Perspectivas para o futuro

Espera-se que os resultados do projeto contribuam para o desenvolvimento de ferramentas de apoio à decisão em políticas públicas de saúde, voltadas à prevenção precoce da obesidade infantil. A longo prazo, o estudo pode subsidiar intervenções personalizadas baseadas em perfis de risco, promovendo ações mais eficazes e sustentáveis na atenção primária à saúde.

“Nossa equipe de epidemiologistas e nossos parceiros estatísticos e da área de matemática computacional temos estabelecido uma eficiente comunicação alinhada a nossos saberes que está resultando ser eficiente e produtiva em nossos primeiros encontros.”

Jorge Melendez
Coordenador do projeto



PARCERIA



FINANCIADORES



COORDENAÇÃO

