

# ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

DISRUPTIVE ARCHITECTURES: IOT, IOB & GENERATIVE IA



**Econo Med<sup>+</sup>**

Medicina Acessível

RM: 98266 / Mel Maia Rodrigues

RM: 98078 / Augusto Barcelos Barros

RM: 98570 / Gabriel Souza de Queiroz

RM: 97707 / Lucas Pinheiro de Melo

## Sumário

ANÁLISE E DESENVOLVIMENTO DE SISTEMAS .....	1
Problemática .....	3
Público-alvo.....	5
Evolução do Projeto .....	7
Pitch .....	13
Organização no GitHub.....	14
Referencias .....	15

## Problemática

A elevada carga de despesas enfrentada pelas operadoras de planos de saúde e seus beneficiários é uma preocupação relevante. Atualmente, o valor do plano de seguro de vida depende predominantemente da faixa etária do beneficiário, resultando em disparidades, com os indivíduos mais jovens pagando menos e os mais idosos pagando mais. A prática em questão reflete a percepção das empresas de saúde de que os indivíduos mais velhos apresentam mais dificuldades de saúde, embora não sejam exclusivamente eles os principais usuários dos serviços de saúde.

Além do valor inicial estabelecido em contrato, há os reajustes anuais, aplicados no aniversário da contratação do plano. Esses reajustes variam de acordo com cada plano e incluem o aumento por sinistralidade, que é quando o número de procedimentos e atendimentos cobertos supera as expectativas da operadora em um determinado período.

A verdadeira razão por trás dos altos custos dos planos de saúde está associada a diversos fatores. A mão de obra na área da medicina é cara, devido aos altos custos de formação e especialização dos profissionais. Além disso, há os gastos com manutenção de equipamentos e reposição de materiais, que também representam um ônus significativo para os hospitais e clínicas.

Esses fatores, somados à complexidade dos procedimentos médicos e aos avanços tecnológicos constantes na área da saúde, contribuem para o encarecimento dos planos de saúde, pois às operadoras custeiam cada consulta realizada pelo seu beneficiário, podendo ser uma consulta mais cara ou mais barata, dependendo do hospital que o beneficiário recorre.

## **Objetivo**

A meta primordial dessa solução é entregar aos clientes um atendimento de saúde que seja não apenas eficaz em termos de custo, mas também personalizado e de alta qualidade, visando a minimização dos gastos para a operadora de saúde. A abordagem proposta se apoia de forma integral em análises avançadas de dados (analytics), as quais permitem que o software realize uma minuciosa caracterização do perfil de cada beneficiário, valendo-se tanto das informações já disponíveis sobre o cliente quanto de dados externos cuidadosamente selecionados. Esse perfilamento detalhado possibilita direcionar o beneficiário para o tipo de atendimento mais vantajoso em termos de custo, encaminhando-o para clínicas ou hospitais especializados na área correspondente à sua queixa de saúde, porém que pratiquem preços mais acessíveis para a operadora de saúde.

No processo de análise do perfil do beneficiário, são levadas em conta uma série de informações relevantes, tais como dados demográficos, histórico completo de utilização dos serviços de saúde, mapeamento dos melhores prestadores na região onde o beneficiário se encontra, consulta ao dicionário de Classificação Internacional de Doenças (CID) para compreensão precisa das condições de saúde do paciente, e ainda uma base de dados abrangente dos prestadores de saúde que possuem expertise no tratamento de comorbidades específicas. A partir desses dados, a operadora é capaz de identificar não apenas o tipo de paciente, mas também suas tendências e necessidades de saúde mais frequentes, permitindo assim a seleção de hospitais ou clínicas próximos que ofereçam um custo por paciente mais baixo, sem comprometer a qualidade do atendimento necessário ao paciente. Esse enfoque estratégico não apenas proporciona economia financeira à operadora, mas também assegura que os beneficiários recebam o cuidado mais adequado às suas condições, resultando em uma experiência de saúde mais satisfatória e eficaz.

## **Público-alvo**

A solução é especificamente voltada para operadoras de planos de saúde, que são empresas dedicadas a oferecer assistência médica e cobertura de saúde para indivíduos e grupos mediante pagamento. Além de comercializar planos, elas fornecem uma ampla gama de serviços e atendimento ao cliente.

Em 2000, a Agência Nacional de Saúde Suplementar (ANS) categorizou as operadoras em sete modalidades:

**Administradora:** Responsável pela gestão administrativa e financeira dos planos de saúde, atuando como intermediária entre clientes e prestadores de serviços de saúde.

**Cooperativa Médica:** Formadas por médicos que se associam para oferecer serviços de saúde aos membros da cooperativa, com uma abordagem centrada no paciente.

**Cooperativa Odontológica:** Similar à cooperativa médica, porém focada em serviços odontológicos, onde os dentistas se unem para fornecer atendimento odontológico compartilhado.

**Autogestão:** Os próprios beneficiários administram e financiam o plano de saúde, tomando decisões sobre cobertura e serviços oferecidos.

**Filantropia:** Operadoras que visam o benefício da comunidade, oferecendo serviços a preços acessíveis ou gratuitos para grupos vulneráveis, muitas vezes com apoio governamental ou por meio de doações.

**Medicina de Grupo:** Oferecem planos de saúde para grupos específicos, como empresas ou associações, onde os custos são compartilhados pelos beneficiários conforme os termos do plano.

**Odontologia de Grupo:** Semelhante à medicina de grupo, mas focada em serviços odontológicos, com cobertura para consultas, exames e tratamentos dentários.

Em 2001, a Lei nº10.185 exigiu que as seguradoras que atuavam no segmento de seguro saúde se transformassem em seguradoras especializadas, sujeitas a uma nova regulação e fiscalização vinculada ao Ministério da Saúde.

No Brasil, destacam-se grandes operadoras como NotreDame Intermédica, Hapvida Assistência Médica, Bradesco Saúde, Amil Assistência Médica, SulAmerica Saúde, Seguros Unimed Saúde, Prevent Sênior, Assim Saúde e Porto Seguro Saúde. Juntas, essas nove maiores operadoras atendem cerca de 20,87 milhões de pessoas, aproximadamente 41,58% dos 50,2 milhões de indivíduos com planos médicos privados no país, conforme dados da ANS em 2023.

A solução uma vez desenvolvida e testada, será disponibilizada às operadoras de saúde de todas as modalidades, exceto as de Filantropia. Tendo um foco maior nas nove maiores operadoras do Brasil, já citadas anteriormente.

## **Evolução do Projeto**

### **Sprint 1: Perfilamento e Preparação dos Dados**

Na primeira sprint, focamos na análise e preparação dos dados utilizando as seguintes bibliotecas:

- NumPy
- Pandas
- Seaborn
- Matplotlib
- Scikit-learn.

Para avaliar a qualidade dos modelos iniciais, aplicamos métricas de desempenho como:

- Erro Quadrático Médio (MSE).
- Erro Absoluto Médio (MAE).
- Coeficiente de Determinação ( $R^2$ ).

Esses recursos nos permitiram explorar o comportamento dos dados e estabelecer uma base sólida para o desenvolvimento do modelo.

### **Sprint 2: Implementação de Redes Neurais**

Na segunda sprint, aprimoramos a capacidade preditiva do sistema com a introdução de redes neurais utilizando TensorFlow.

Arquitetura da Rede Neural:

- Modelo sequencial composto por camadas densas.
- Função de ativação ReLU para capturar relações não lineares entre as variáveis.
- Camadas ocultas com 428, 364, 166 e 66 neurônios.
- Camada de saída com 1 neurônio para prever os custos.

Essa abordagem proporcionou maior flexibilidade, capturando relações mais complexas entre as variáveis e melhorando a performance preditiva.

### **Sprint 3: API para Previsão de Custos de Saúde**

Na terceira sprint, desenvolvemos uma API para tornar o modelo preditivo acessível e utilizável em tempo real.

A API, criada com FastAPI, é capaz de fornecer previsões de custos de saúde para pacientes com base em informações demográficas e hábitos de vida, utilizando o modelo de rede neural treinado na segunda sprint.

#### **Detalhes da API**

- Endpoint: /predict/
- Método: POST

#### **Descrição:**

Este endpoint recebe os dados de um paciente segurado e retorna a previsão dos custos de saúde com base no modelo de rede neural.

#### **Parâmetros da Requisição:**

Os dados devem ser enviados no formato JSON, contendo as seguintes informações:

- age (int): Idade do paciente.
- sex (str): Gênero do paciente ("male" ou "female").
- bmi (float): Índice de Massa Corporal (BMI).
- children (int): Número de filhos.
- smoker (str): Status de fumante do paciente ("yes" ou "no").
- region (str): Região de residência do paciente ("northeast", "northwest", "southeast", ou "southwest").



## **Frameworks**

Nesse momento, focamos em desenvolver primeiramente um dos focos, o perfilamento do paciente.

### **Análise de dados**

Nos treinamentos realizados, foram utilizados principalmente dois frameworks para aprendizado de máquina e redes neurais:

#### **Scikit-learn (sklearn):**

##### **O que é?**

O Scikit-learn é uma biblioteca em Python que oferece uma ampla variedade de algoritmos de aprendizado de máquina supervisionados e não supervisionados, além de ferramentas para pré-processamento de dados, validação de modelo e métricas de avaliação

- **Modelagem de Regressão Linear**

O Scikit-learn está sendo utilizado para construir e treinar o modelo de regressão linear, permitindo prever o valor do convenio para os pacientes com base em variáveis diferentes.

- **Pré-processamento de Dados:**

Ferramentas como `train_test_split` serão usadas para dividir os dados em conjuntos de treino e teste, enquanto `StandardScaler` será aplicado para padronizar os dados, garantindo que diferentes características tenham escalas comparáveis.

- **Avaliação de Modelos**

Métodos como Mean Squared Error (MSE), Mean Absolute Error (MAE) e R-squared ( $R^2$ ) Score serão empregados para avaliar a qualidade dos modelos de regressão linear desenvolvidos.

## **TensorFlow:**

### **O que é?**

TensorFlow é uma biblioteca de código aberto para aprendizado de máquina e inteligência artificial desenvolvida pela Google Brain Team. Ela oferece uma série de recursos para criar e treinar modelos de aprendizado profundo.

- **Construção de Redes Neurais**

O TensorFlow está sendo utilizado para criar e treinar redes neurais profundas. Isso permite uma modelagem mais complexa dos dados, capturando relações não-lineares e interações entre as variáveis.

- **Compilação e Treinamento de Modelos**

Com o TensorFlow, os modelos de rede neural são compilados com otimizadores e funções de perda apropriados, como o Adam e o Mean Squared Error (MSE), respectivamente. Em seguida, os modelos são treinados utilizando o método fit, onde são ajustados aos dados de treinamento ao longo de várias iterações (epochs).

- **Avaliação de Modelos de Rede Neural**

Após o treinamento, os modelos de rede neural estão sendo avaliados em dados de teste para determinar sua eficácia. Métricas como o erro médio absoluto (MAE) podem ser utilizadas para medir a precisão das previsões feitas pelo modelo.

Esses frameworks desempenham papéis distintos nos treinamentos:

**Scikit-learn:** Foi utilizado principalmente para treinamento de modelos de regressão linear, manipulação de dados e avaliação de desempenho.

**TensorFlow:** Foi utilizado para construir e treinar redes neurais profundas, permitindo uma modelagem mais complexa e captura de relações não-lineares nos dados.

### **Front-end**

No front-end será utilizado o react-native, pois o software será desenvolvido para dispositivos mobiles.

### **Back-end**

O back-end será desenvolvido na linguagem Java, usando o framework Spring Boot, para auxiliar o desenvolvedor.

### **Banco De Dados**

Como haverá a análise de muitos dados, o banco relacional está sendo o ideal para a aplicação nessa primeira etapa.

## **Machine Learning (ML) e Inteligência Artificial (IA)**

No projeto, atualmente, os conceitos de Machine Learning (ML) e Inteligência Artificial (IA), estão sendo aplicados de inúmeras maneiras para alcançar o objetivo, os conceitos estão sendo utilizados como:

### **Modelagem de Regressão Linear:**

- Utilizando técnicas de ML, estamos construindo modelos de regressão linear para prever o valor do convênio para os pacientes com base em diferentes variáveis, como idade, gênero, hábitos de fumar e região geográfica.
- Esses modelos são treinados com dados históricos para aprender **padrões e relações entre os recursos e o valor do convênio, permitindo fazer previsões precisas para novos pacientes.**

### **Redes Neurais:**

- Além da regressão linear, estamos explorando o poder das redes neurais profundas, para modelar relações mais complexas e não-lineares nos dados.
- Essas redes são compostas por várias camadas de neurônios que aprendem representações abstratas dos dados, permitindo capturar sutilidades e interações mais dentre as variáveis.

### **Pré-processamento de Dados e Avaliação de Modelos:**

- Antes de alimentar os dados aos modelos de ML e IA, é realizada uma etapa de pré-processamento, que inclui limpeza, transformação e padronização dos dados.
- Após o treinamento dos modelos, são utilizadas métricas de avaliação, como MSE, MAE e  $R^2$  Score, para medir o desempenho e a precisão dos modelos em fazer previsões.

## **Pitch**

Sprint 2: <https://youtu.be/xmCTCtY1BFI>

Sprint 3: <https://youtu.be/RBsQWQ6p4vw>

## **Organização no GitHub**

<https://github.com/FiapChallenge>

## Referencias

<https://www.beneficiosrh.com.br/maiores-operadoras-planos-de-saude/>

<https://www.portaldaindustria.com.br/industria-de-a-z/saude-suplementar-o-que-e-e-como->

<funciona/#:~:text=Alguns%20exemplos%20de%20operadoras%20de,SulAm%C3%A9rica%2C%20Unimed%2C%20entre%20outras.>

<https://fenasaude.org.br/conteudos/o-que-e-saude-suplementar>

<https://www.piwi.com.br/blog/entenda-o-que-e-uma-operadora-de-plano-de-saude/#:~:text=cada%20uma%20delas.->

<,Como%20funciona%20uma%20operadora%20de%20plano%20de%20sa%C3%BAde%3F,pagamento%20%C3%A0%20Administradora%20do%20plano.>

<https://idec.org.br/consultas/dicas-e-direitos/conheca-os-tipos-de-reajuste-possiveis/#:~:text=H%C3%A1%20tr%C3%AAs%20modalidades%20diferentes%20de,situa%C3%A7%C3%B5es%20eles%20podem%20ser%20aplicados.>

<https://agenciabrasil.ebc.com.br/saude/noticia/2024-01/numero-de-usuarios-de-planos-de-saude-passa-de-51-milhoes-em-2023#:~:text=O%20setor%20engloba%20680%20operadoras,rela%C3%A7%C3%A3o%20a%20dezembro%20de%202022.>

<https://www.gov.br/ans/pt-br/assuntos/noticias/numeros-do-setor/ans-divulgados-economico-financeiros-relativos-ao-2o-trimestre-de-2023>