



## Controle do Documento

### Histórico de revisões

Data	Autor	Versão	Resumo da atividade
01/02/2023	Yago Phellipe Matos Lopes	0.1	Preenchimento da seção 4.1.3.
02/02/2023	José Vitor Marcelino	0.2	Implementação de Personas 4.1.6
06/02/2023	José Vitor Marcelino	0.3	Rascunho da Introdução 1.0 (Pré-Validação com a professora)
06/02/2023	Emely Vitória	0.4	Preenchimento da seção 4.1.1
07/02/2023	Marcos Teixeira	0.5	Preenchimento da seção 4.1.5
08/02/2023	Marcos Teixeira	0.6	Preenchimento da seção 4.1.5
08/02/2023	Yuri Toledo	0.7	Preenchimento da seção 4.1.7
09/02/2023	Marcos Teixeira	0.8	Preenchimento da seção 4.1.3
10/02/2023	Emely Vitória	0.9	Preenchimento da seção 2.1

# Sumário

<b>1. Introdução</b>	<b>4</b>
<b>2. Objetivos e Justificativa</b>	<b>5</b>
2.1. Objetivos	5
2.2. Proposta de Solução	5
2.3. Justificativa	5
<b>3. Metodologia</b>	<b>6</b>
<b>4. Desenvolvimento e Resultados</b>	<b>7</b>
4.1. Compreensão do Problema	7
4.1.1. Contexto da indústria	7
4.1.2. Análise SWOT	7
4.1.3. Planejamento Geral da Solução	7
4.1.4. Value Proposition Canvas	7
4.1.5. Matriz de Riscos	7
4.1.6. Personas	8
4.1.7. Jornadas do Usuário	8
4.2. Compreensão dos Dados	9
4.3. Preparação dos Dados e Modelagem	10
4.4. Comparação de Modelos	11
4.5. Avaliação	12
<b>5. Conclusões e Recomendações</b>	<b>13</b>
<b>6. Referências</b>	<b>14</b>
<b>Anexos</b>	<b>15</b>

# 1. Introdução

*Apresente de forma sucinta o parceiro de negócio, seu porte, local, área de atuação e posicionamento no mercado. Maiores detalhes deverão ser descritos na seção 4. Descreva resumidamente o problema a ser resolvido (sem ainda mencionar a solução).*

O ICESP (Instituto do Câncer de São Paulo) é um dos maiores hospitais especializados no tratamento de câncer da América Latina. Sua sede é um prédio com mais de 100 metros de altura e 80 mil metros quadrados, localizado na Avenida Doutor Arnaldo, na zona oeste da cidade de São Paulo.

O instituto surgiu em 2008 (Embora o edifício já exista há mais de 30 anos) como resultado de uma parceria entre a Faculdade de Medicina da USP (FMUSP) e o governo de São Paulo. No começo da elaboração de seu projeto (1987), o instituto inicialmente tinha a finalidade de ser um centro médico com programas focados na área da saúde da mulher. Anos antes de sua inauguração, em 2005, o Conselho Deliberativo do HCFMUSP (Hospital das Clínicas da Faculdade de Medicina da USP) apresentou uma proposta que tornaria o projeto não só uma unidade exclusiva para saúde feminina, mas também um hospital que realizaria transplantes, cirurgias complexas e oncologia cirúrgica. Ao contrário do previsto, em dezembro de 2007, por uma decisão do governador de São Paulo da época, o hospital se tornou uma unidade totalmente dedicada ao paciente oncológico.

O instituto tem como principal área de atuação o tratamento de câncer, já que seu principal propósito é ser referência de assistência, ensino e pesquisa em oncologia. Inclusive dentro do ICESP existe o CTO (Centro de Investigação Translacional em Oncologia), que possui vários grupos de pesquisadores que investigam novos medicamentos e métodos de tratamento.

Por ser uma organização pública, o ICESP não possui rivais de mercado, já que seus serviços possuem um viés colaborativo/educativo e suas pesquisas podem ser utilizadas por outras instituições da área que estejam interessadas em contribuir com a ciência e a sociedade. Existem outros hospitais que também trabalham na área da oncologia e todos (incluindo o ICESP) estão investindo bastante em tecnologia e em processos inovadores (como por exemplo o uso das análises diagnóstica, preditiva e prescritiva de dados).

Devido ao seu foco em oncologia, o ICESP enxerga que a evolução do câncer de mama ainda é algo bastante variável durante os tratamentos oncológicos convencionais, sendo assim,

o Instituto do Câncer do Estado de São Paulo deseja descobrir um padrão preditivo existente entre os pacientes diagnosticados com câncer para saber qual tipo de tratamento é o melhor em cada caso: Neoadjuvante (1º quimioterapia e 2º cirurgia) ou Adjuvante (1º cirurgia e 2º quimioterapia).

## 2. Objetivos e Justificativa

### 2.1. Objetivos

O principal objetivo do parceiro de negócio é viabilizar a criação de uma solução capaz de:

- Analisar uma ampla variedade de informações, incluindo dados médicos (informações sobre saúde, histórico médico e medicações) e demografia (informações sobre idade, gênero, renda, educação, etc).
- Filtrar as informações essenciais do paciente para prever uma saída futura com base em dados e relações entre as variáveis de entrada.
- Classificar dados de pacientes com câncer a fim de identificar qual é a melhor forma de realizar o tratamento de câncer de mama: neo (1º quimioterapia e 2º cirurgia) ou adjuvante (1º cirurgia e 2º terapia), para, assim, detectar padrões que indiquem aos profissionais de saúde uma possível rota de tratamento indicado para cada perfil dos pacientes.

### 2.2. Proposta de Solução

A equipe desenvolverá um modelo preditivo que tem como objetivo recomendar o tratamento adequado para cada caso específico. O produto vai focar em resolver o problema da variabilidade de respostas a tratamentos de câncer de mama, prevendo qual dos dois tratamentos (adjuvante e neoadjuvante) é o melhor para cada paciente. Dessa forma, auxiliando o médico na escolha do tratamento apropriado. O modelo terá os dados de pacientes diagnosticados com câncer fornecidos pelo parceiro como base, esses dados servirão para análise e treinamento do algoritmo do modelo.

### 2.3. Justificativa

O tratamento convencional do câncer de mama possui resultados muito variados, o que certamente atrapalha qualquer conclusão médica na hora da sugestão de qual tratamento deve ser o adequado. A ordem dos processos médicos (quimioterapia e cirurgia) aplicados durante o tratamento de um paciente tem influência direta na taxa de sucesso de remoção do câncer, sendo assim, muito importante definir qual ordem é a mais apropriada a seguir. Tendo esse problema em vista, um modelo preditivo baseado na análise de dados clínico-laboratoriais é uma solução que pode auxiliar o médico e sustentar, por meio de ferramentas, qual tratamento (adjuvante ou neoadjuvante) é o melhor para o caso de cada paciente.

## 3. Metodologia

*Descreva as etapas da metodologia CRISP-DM que foram utilizadas para o desenvolvimento, citando o referencial teórico. Você deve apenas enunciar os métodos, sem dizer ainda como ele foi aplicado e quais resultados obtidos.*

## 4. Desenvolvimento e Resultados

### 4.1. Compreensão do Problema

#### 4.1.1. Contexto da indústria

Após uma análise de mercado, não foram identificadas empresas que possam ser consideradas concorrentes diretas do parceiro de negócios, já que ele se posiciona como uma instituição pública dedicada a servir a sociedade. Todo o conhecimento produzido internamente pode ser usado por outras organizações, desde que sejam respeitadas as regras de uso e direitos comerciais. No entanto, existem outras empresas que atuam na mesma área, como o Hospital A.C. Camargo, especializado no tratamento e pesquisa do câncer; O Centro oncológico Família Dayan - Daycoval, pertence ao Hospital Israelita Albert Einstein, e o INCA, um órgão vinculado ao Ministério da Saúde responsável pela prevenção e controle do câncer no Brasil.

O Hospital das Clínicas é uma instituição de saúde pública universitária, localizada em São Paulo, Brasil, que possui um setor interno para o tratamento de CA, o Instituto do Câncer do Estado de São Paulo (ICESP). Como uma instituição pública, o Hospital não tem um modelo de negócios tradicional baseado em lucro, pois é financiado pelo governo e pela universidade para fornecer atendimento médico gratuito à população. Além disso, o Hospital das Clínicas também oferece serviços médicos pagos, como exames e procedimentos, para aqueles que têm capacidade financeira para pagá-los, o que ajuda a complementar o financiamento da instituição. Em resumo, o modelo de negócios do Hospital das Clínicas é misto, incluindo financiamento público e serviços médicos pagos para complementar sua operação.

Considerando o ambiente de mercado em que o hospital está inserido, pode-se ver que ele está sempre em busca de melhorias, investindo em modernização, inovação tecnológica e plataforma educacional, além de firmar parcerias com a iniciativa privada e expandir o processo de internacionalização para ser um centro educacional de referência no mundo (USP, 2021). Essas inovações incluem a procura por equipamentos mais avançados, tratamentos mais eficazes e menos prejudiciais, entre outros aspectos.

Abaixo uma análise de indústria utilizando as cinco forças de Porter.

**Rivalidade entre os concorrentes:** por ser uma instituição governamental sem fins lucrativos, o Instituto do Câncer de São Paulo não tem concorrentes diretos, mas sim hospitais parceiros com os quais trabalha na pesquisa e no tratamento do CA. No entanto, existem outras instituições, como o Instituto Nacional do Câncer, e vários hospitais privados que também oferecem tratamento para o câncer e competem por pacientes.



**Poder de barganha de clientes:** é variável, dependendo do nível socioeconômico e da disponibilidade de opções de saúde para eles. No entanto, a instituição possui uma grande demanda por parte da população, o que pode limitar o poder de negociação dos clientes.

**Poder de barganha de fornecedores:** o parceiro precisa de produtos muito avançados e de elevado valor agregado, o que significa que seus fornecedores têm grande poder de barganha. Isso acontece porque certos medicamentos e equipamentos de pesquisa não são amplamente disponíveis e sua produção e preço são controlados por um pequeno grupo de empresas. Como fornecedores de dados, que são a base para as atividades do parceiro, os hospitais também podem recusar fornecê-los. Além disso, as tecnologias criadas requerem um longo processo de licenciamento, e os responsáveis pelo desenvolvimento podem fazer exigências quanto ao custo da solução.

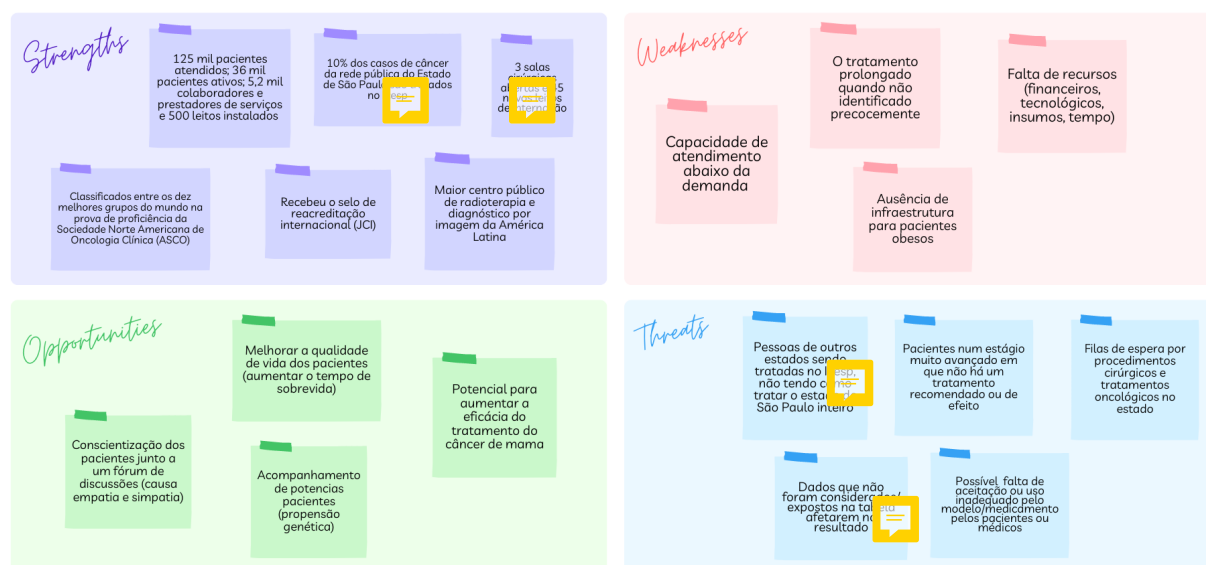
**Ameaças de produtos substitutos:** é moderado devido ao fato de que há cada vez mais investimentos da iniciativa privada no desenvolvimento de novas tecnologias e serviços, podendo haver novas soluções e tratamentos mais eficazes.

**Ameaças de novos entrantes:** não há preocupação com a chegada de novos competidores, pois as instituições na mesma área de atuação do Hospital das Clínicas não são vistas como concorrentes, mas sim como parceiros. A FMUSP é reconhecida como uma referência na área da saúde, e qualquer nova empresa que surja no mercado não seria capaz de competir com ela a curto prazo.

## 4.1.2. Análise SWOT

Figura 1 - Matriz SWOT

### SWOT Analysis – Instituto do Câncer do Estado de São Paulo



Fonte: elaboração própria

#### (Forças)

De acordo com a plataforma oficial do Instituto do Câncer do Estado de São Paulo (ICESP), - é responsável por cerca de 10% dos casos de câncer da rede pública do Estado de São Paulo - desde a sua inauguração em 6 de maio de 2008, o instituto atendeu aproximadamente 125 mil pacientes, com 36 mil pacientes ativos e contam com 5,2 mil colaboradores e prestadores de serviços. Além disso, há 445 leitos operacionais e na sua capacidade máxima disponibilizará 490 leitos, sendo 85 de Unidade de Terapia Intensiva (UTI). - <https://icesp.org.br/>

Recentemente, houve uma ampliação em sua infraestrutura onde foram abertas mais três salas cirúrgicas e 45 novos leitos de internação, sendo 15 são de UTI. Isso possibilitará o atendimento de 1.250 novos pacientes no próximo ano e a realização de 840 cirurgias adicionais, o que seria o equivalente a um incremento de 20% do previsto para o período. A iniciativa visa reduzir a fila atual de pacientes oncológicos no Estado de São Paulo em 40%, nos primeiros meses da ação. -

<https://icesp.org.br/noticias/residentes-da-oncologia-clinica-do-icesp-obtem-media-mais-alta-em-exame-mundial/>

O Icesp foi o primeiro hospital da rede pública da capital a receber o selo de reacreditação pela Joint Commission International (JCI) - uma certificação internacional que reconhece a excelência no atendimento e serviços oferecidos à população - em 2014. O selo passa por um processo de verificação que ocorre de três em três anos, o instituto se submete à avaliação a fim de manter o selo, a última verificação ocorreu em 2020.

<https://www.saopaulo.sp.gov.br/spnoticias/orgaos-governamentais/secretaria-da-saude/instituto-do-cancer-de-sao-paulo-recebe-selo-de-reacreditacao-internacional/>

Sobre a qualidade do ensino e pesquisa realizados, o Icesp se destaca no Programa de Residência Médica em Cancerologia Clínica, originado em 1998 como primeiro do país. Desde sua criação, já foram formados mais de 170 médicos oncologistas e atualmente é um dos maiores programas do país, reconhecido nacional e internacionalmente.

Nos últimos anos, os formandos foram sistematicamente classificados entre os dez melhores grupos do mundo na prova de proficiência da Sociedade Norte Americana de Oncologia Clínica (ASCO). Porém, no ano de 2022, os residentes do segundo e terceiro ano do programa de residência médica de Oncologia Clínica do Icesp, alcançaram o melhor desempenho no exame anual da Sociedade Americana de Oncologia Clínica (ASCO). O grupo obteve a maior média entre todas as instituições avaliadas e atingiu sua maior pontuação em relação a avaliações de anos anteriores, superando assim sua melhor marca. O que coloca a instituição no topo dos melhores profissionais de oncologia do mundo. -

<https://icesp.org.br/noticias/residentes-da-oncologia-clinica-do-icesp-obtem-media-mais-alta-em-exame-mundial/>

(Fraquezas)

Um grande agravante na falha do processo do tratamento e na alta demanda, seria a falta de recursos. Por exemplo, a imunoterapia pode trazer benefícios como a alta eficácia e baixos efeitos colaterais, no entanto, uma única caixa desse medicamento pode custar 15 mil reais. Contudo, no Sistema Único de Saúde (SUS) são poucas as novas drogas que são oferecidas, essa dificuldade de acesso frequentemente leva a algumas pessoas a até entrarem na justiça para a obtenção do tratamento.

Segundo Maria Del Pilar Estevez Diz, diretora do Corpo Clínico do Instituto do Câncer do Estado de São Paulo, diz que a dificuldade de acesso ao uso de novas tecnologias é uma das principais razões para que o câncer de mama tenha ainda alta mortalidade no Brasil. "A maior parte dos médicos atua no SUS e no setor privado e, com isso, vivencia situações muito díspares. Falta equidade." Como há uma alta demanda pelos tratamentos, os médicos não têm

tempo suficiente para propriamente entender o paciente num nível mais pessoal, o que faz com que a atenção e o atendimento não sejam tão precisos.

O oncologista Stephen Stefani, presidente da International Society for Pharma-coeconomics and Outcomes research (Ispor) no Brasil, relata que uma das origens do problema de acesso são as distorções no sistema. “Apenas 25% das pessoas no Brasil têm acesso a planos de saúde, mas 55% dos recursos no País são gastos com essa população”, afirmou. Além disso, há novos tratamentos contra o câncer que podem custar até US\$ 10 mil quando chegam ao mercado – e pesam no sistema, aumentando a distorção. “Qualquer incorporação de um novo medicamento, se não for feita com cuidado, pode aumentar o número de excluídos. Os recursos são limitados, e não podemos conceder qualquer tipo de desperdício.”

<https://www.anahp.com.br/noticias/aceso-a-novos-tratamentos-pelo-sus-ainda-e-um-obstaculo/>

(Ameaças)

O Ministério Público Federal (MPF) entrou com uma ação contra o governo de São Paulo para que providências sejam tomadas a respeito da lei federal que determina que pacientes com câncer devem receber tratamento em até 60 dias após o diagnóstico. O MPF destacou que mais de 18 mil pessoas aguardam mais de dois meses entre o diagnóstico e o começo da terapia.

De acordo com a Secretaria Estadual da Saúde (SES), 1.536 pessoas continuam na espera por cirurgias para tratamento de câncer no estado e, em alguns casos, a espera chega a ser de oito meses.

<https://g1.globo.com/sp/sao-paulo/noticia/2023/01/24/com-mais-de-mil-pacientes-com-cancer-a-espera-de-cirurgia-governo-de-sp-anuncia-45-leitos-e-3-salas-cirurgicas-na-tentativa-de-reduzir-fila.ghtml>

Para poder ser atendido no Icesp, o paciente precisa ser de uma certa localidade para poder ser atendido, mas há casos em que alguns pacientes trocam o seu comprovante de residência para poder receber o tratamento.

A boa alimentação é um fator importante para a prevenção do câncer, manter uma dieta equilibrada pode ajudar na prevenção de diversas doenças. Alimentos ricos em fibras, vitaminas e antioxidantes oferecem inúmeros benefícios ao organismo. Portanto, a má alimentação da população pode agravar os casos de câncer.

<https://icesp.org.br/mitos-e-verdades-sobre-o-cancer/>

(Oportunidades)

A Organização Mundial da Saúde (OMS) recomenda ações de prevenção, detecção precoce e acesso ao tratamento para controle melhor do câncer, pois quanto mais cedo o câncer for identificado, maiores são as chances de cura. A detecção precoce do câncer consiste em duas estratégias: a primeira seria o rastreamento, que tem por objetivo encontrar o câncer pré-clínico ou as lesões pré-cancerígenas, por meio de exames de rotina em uma população-alvo sem sinais e sintomas sugestivos do câncer rastreado. A segunda, corresponde ao diagnóstico precoce, que busca identificar o câncer em estágio inicial em pessoas que apresentam sinais e sintomas suspeitos da doença.

<https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/deteccao-precoce-do-cancer.pdf>

A conscientização da população também é um fator importante para a prevenção do câncer. Por exemplo, segundo estatísticas do Instituto Nacional de Câncer (Inca) o tabagismo é a principal causa de câncer de pulmão evitável no mundo e as consequências da queima do cigarro são sentidas não apenas por quem fuma, mas também por todos ao seu redor. Outro exemplo seria o consumo de álcool que, de acordo com a Agência Internacional de Pesquisa sobre o Câncer, a quantia de 18 gramas (aproximadamente duas doses) de álcool por dia era suficiente para aumentar significativamente o risco de desenvolver câncer de mama. Consequentemente, com pessoas mais bem conscientizadas e proativas sobre o assunto, o número de pacientes diminui.


[https://www.hcor.com.br/imprensa/noticias/oncologista-do-hcor-aponta-10-dicas-para-prevencao-do-cancer/?gclid=CjwKCAiAioifBhAXEiwApzCztpKeXJbn6tunOQIO8T6Cawb40AZJ6SFccPqH2riiD\\_Gx1Moi2MEvoBoCQoUQAvD\\_BwE](https://www.hcor.com.br/imprensa/noticias/oncologista-do-hcor-aponta-10-dicas-para-prevencao-do-cancer/?gclid=CjwKCAiAioifBhAXEiwApzCztpKeXJbn6tunOQIO8T6Cawb40AZJ6SFccPqH2riiD_Gx1Moi2MEvoBoCQoUQAvD_BwE)



### 4.1.3. Planejamento Geral da Solução

#### a) Qual é o problema a ser resolvido?

A maior dificuldade que médicos de câncer de mama enfrentam atualmente consiste em decidir qual é o tratamento ideal para cada paciente, posto que o método varia entre cada indivíduo. Dessa forma, desprovidos de uma métrica acurada, os médicos tendem a gastar muito tempo analisando os dados de cada paciente para tentar ver o padrão daquela pessoa que, por ser um processo manual, ainda está sujeito a erros, podendo ser indicado o tratamento que não seria tão eficaz para aquele indivíduo. E, desse modo, precisar de um modelo preditivo para

responder com certeza aos questionamentos: "Realizar a cirurgia primeiro e a quimioterapia depois?"; "Realizar a quimioterapia primeiro  a cirurgia depois?".

#### **b) Qual a solução proposta (Visão de negócios).**

Desenvolver uma inteligência artificial para ajudar médicos a escolher o melhor tratamento para seus pacientes com câncer de mama. A plataforma coleta e analisa múltiplos dados de pacientes passados, incluindo informações clínicas e resultados de tratamentos, para criar um modelo preditivo que indica a probabilidade de sucesso de cada tratamento para um paciente específico.

Esta solução resolve o dilema que muitos médicos enfrentam ao tentar decidir qual tratamento é melhor para cada paciente, pois fornece uma base sólida de dados e análise para apoiar suas decisões clínicas. Além disso, a plataforma pode ajudar a garantir que os pacientes recebam o tratamento mais eficaz e aumentar a eficiência do sistema de saúde, pois permite que os médicos tomem decisões mais informadas e baseadas em evidências.

A partir de uma perspectiva de negócios, esta solução pode se destacar em um mercado em constante evolução e com crescente demanda por soluções tecnológicas avançadas na medicina. Além disso, a plataforma pode ser oferecida como um serviço a hospitais, clínicas e grupos médicos, dispostos a pagar pela realização do serviço, gerando um grande lucro.

#### **c) Qual o tipo de tarefa (regressão ou classificação).**

Classificação, pois o output não é contínuo. 

#### **d) Como a solução proposta deverá ser utilizada.**



A solução proposta deverá ser utilizada por um médico especialista no tratamento do paciente, ciente de que o modelo preditivo trata-se de, no máximo, uma *recomendação*. A decisão final sobre o melhor tratamento deve invariavelmente basear-se na análise dos dados coletados pelo médico sobre determinado paciente.

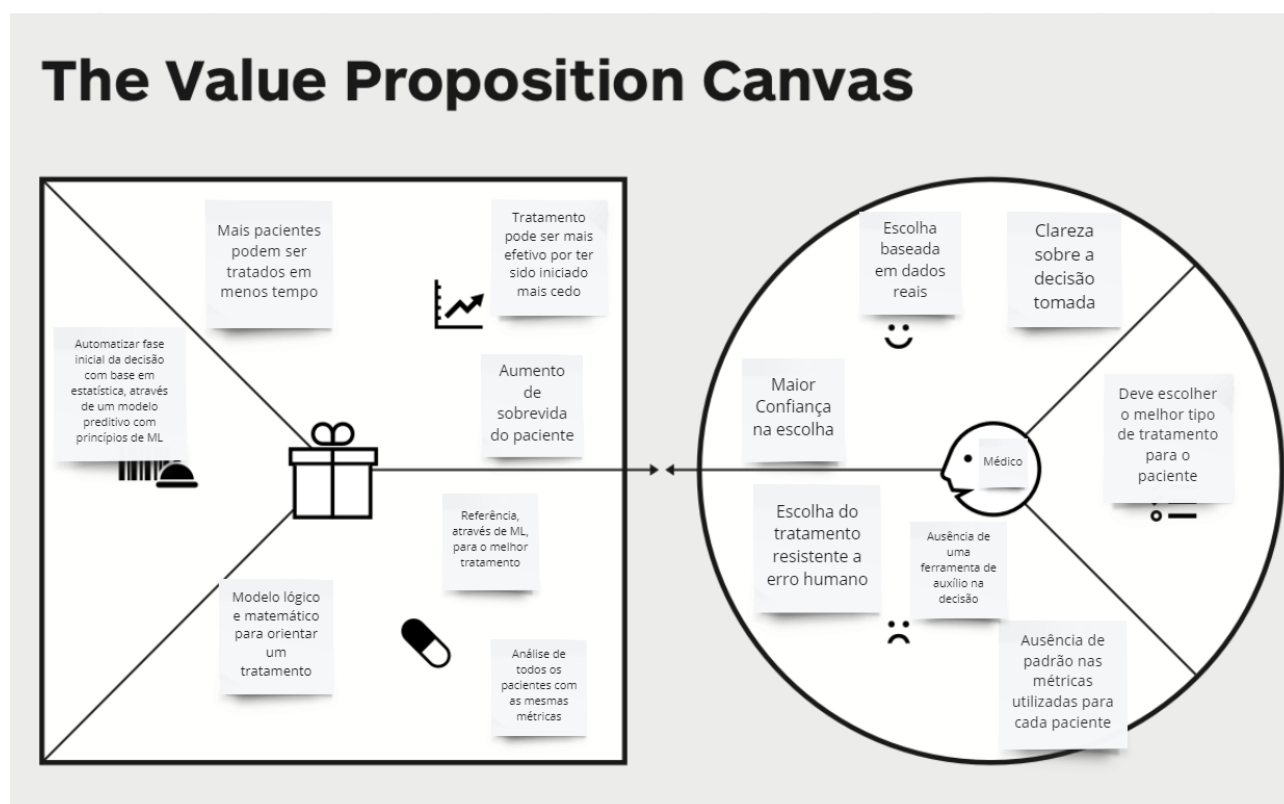
#### **e) Quais os benefícios trazidos pela solução proposta.**

Os benefícios trazidos pela solução proposta consistem na drástica redução do tempo necessário para estabelecer qual é o tratamento mais adequado para cada paciente. Dessa forma, ganha-se em eficiência e até mesmo em efetividade, ainda que de forma incipiente.

#### f) Qual será o critério de sucesso e qual métrica será utilizada para avaliá-lo.

O critério de sucesso é o aumento do tempo de sobrevida dos pacientes, tendo como métricas o fato de a paciente ser reincidente. Se houve extensão no tempo estimado de sobrevida. Ressaltando-se que a extensão do tempo de sobrevida é o principal fator analisado no entendimento do sucesso do modelo.

#### 4.1.4. Value Proposition Canvas



### 4.1.5. Matriz de Risco

A importância da matriz de risco para o nosso projeto ajuda a identificar, avaliar e priorizar os riscos potenciais. Isso nos permite ter uma visão clara e objetiva dos desafios e ameaças enfrentados, bem como das ações a serem tomadas para minimizar seu impacto. Além disso, a matriz de risco também fornece uma base para a monitorização contínua dos riscos e para a atualização dos planos de mitigação, garantindo assim a adaptação a mudanças no ambiente do projeto. Em resumo, a tabela de matriz de risco é fundamental para garantir o sucesso do projeto, tomando medidas preventivas para mitigar ameaças e maximizar oportunidades.

Figura 1 - Matriz de Risco

Probabilidade	Ameaças					Oportunidades					Probabilidade
90%											90%
70%				F			I				70%
50%				G							50%
30%		A	D		H						30%
10%			B	C	E						10%
	Muito Baixo	Baixo	Moderado	Alto	Muito Alto	Muito Alto	Alto	Moderado	Baixo	Muito Baixo	
Impacto											

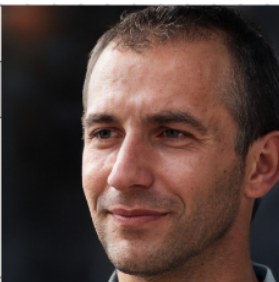
	NOME	CATEGORIA	PROBABILIDADE	IMPACTO
A	Descompromisso com o horário de desenvolvimento do projeto	Desenvolvimento	30%	BAIXO
B	Desentendimento entre os membros da equipe	Comunicação	10%	MODERADO
C	Médicos não fazerem a utilização do modelo preditivo	Desenvolvimento	10%	ALTO




D	Falta de dados complementares	Comunicação	30%	MODERADO
E	Falta de dados essenciais	Comunicação	10%	MUITO ALTO
F	Efeitos colaterais do tratamento	Desenvolvimento	70%	ALTO
G	Abandono do tratamento pelo paciente	Desenvolvimento	50%	ALTO
H	Os tratamentos não fazerem efeitos	Desenvolvimento	30%	MUITO ALTO
I	Ajudar na escolha do melhor tratamento para a paciente diagnosticada com câncer de mama através do modelo preditivo	Desenvolvimento	70%	ALTO

Fonte: elaboração própria

## 4.1.6. Personas

<b>NOME:</b>	Marcos Fernandes Neto	
<b>INFORMAÇÕES PESSOAIS:</b>		
<ul style="list-style-type: none"> <li>Marcos possui 52 anos.</li> <li>Formado em medicina.</li> <li>Trabalha como médico há 22 anos.</li> <li>É casado e possui 2 filhos.</li> <li>Nasceu e mora em São Paulo.</li> </ul>		
<b>DORES:</b>	<ul style="list-style-type: none"> <li>Não sabe para qual tipo de tratamento encaminhar seus pacientes.</li> <li>Às vezes seus pacientes não voltam para dar continuidade ao tratamento.</li> <li>Muitos pacientes só procuram atendimento médico quando o câncer já está em um estágio avançado.</li> </ul>	
<b>OBJETIVOS/NECESSIDADES:</b>		
<ul style="list-style-type: none"> <li>Diminuir a quantidade de casos que necessitam tratamentos mais severos e invasivos.</li> <li>Um modelo preditivo para que ele saiba para qual tratamento encaminhar o paciente.</li> <li>Deseja saber qual é o tratamento que trará mais resultados aos pacientes.</li> </ul>		

NOME:	Renata Gonçalves Dias	
INFORMAÇÕES PESSOAIS:		
<ul style="list-style-type: none"><li>• Renata possui 41 anos.</li><li>• Está fazendo mestrado em medicina.</li><li>• Trabalha como pesquisadora.</li><li>• É casada.</li><li>• Nasceu na Bahia e mora em São Paulo.</li></ul>		
DORES:	<ul style="list-style-type: none"><li>• Não sabe porque alguns pacientes respondem melhor ao tratamento neo enquanto outros respondem melhor ao tratamento adjuvante.</li><li>• Os dados frequentemente possuem informações vazias ou são insuficientes.</li><li>• Desconhece o que influencia na taxa de sucesso dos tratamentos.</li></ul>	
OBJETIVOS/NECESSIDADES:		
<ul style="list-style-type: none"><li>• Saber quais fatores fazem o paciente responder melhor a determinado tratamento.</li><li>• Explorar meios que viabilizem uma diminuição dos efeitos colaterais causados pelos tratamentos.</li><li>• Deseja descobrir se a alteração na sequência dos processos causa uma diminuição no tempo total de tratamento do paciente.</li></ul>		

Ambas personas utilizarão o modelo e serão afetadas por ele no sentido laboral, de maneira em que os resultados desse modelo podem afetar diretamente em seus trabalhos, já que pode mudar completamente a visão desses profissionais (médico e pesquisadora) a respeito do tratamento do câncer de mama.

#### 4.1.7. Jornadas do Usuário

O Mapa de Jornada do Usuário é uma ferramenta que ajuda a entender e acompanhar as fases pelas quais um usuário passa ao interagir com um modelo, no caso específico, médicos que utilizam uma plataforma para realizar análises e predições sobre seus pacientes.

A primeira fase é o Conhecimento da Plataforma. Nesta etapa, o médico entra em contato com o modelo pela primeira vez e precisa compreender o objetivo e o funcionamento da plataforma para que possa utilizá-la da melhor forma possível. É importante que ele saiba manusear a plataforma e entenda suas funcionalidades.

Na segunda fase, a Entrada de Dados, o médico precisa inserir informações sobre a paciente, como dados clínicos, exames, histórico médico, etc., para que o modelo possa fazer a análise e gerar uma predição. É importante que o médico preste atenção aos dados inseridos para garantir a precisão da análise.

Na terceira fase, a Saída de Informações, o médico tem acesso ao resultado gerado pelo modelo, ou seja, à predição. É importante que ele entenda o porquê da predição e que possa interpretar corretamente as informações geradas.

Por fim, a quarta e última fase é o Final do Tratamento, momento em que o médico informa a conclusão do procedimento, fornecendo dados para o modelo aprender e aumentar sua acurácia. É importante que ele preste atenção aos resultados finais para que possa contribuir para o desenvolvimento da plataforma.

Em resumo, o Mapa de Jornada do Usuário é uma ferramenta valiosa para entender e acompanhar o processo de interação do médico, nosso principal usuário, com a plataforma, garantindo que ele possa utilizá-la de maneira eficiente e eficaz.

Figura 2 - Mapa de Jornada de Usuário

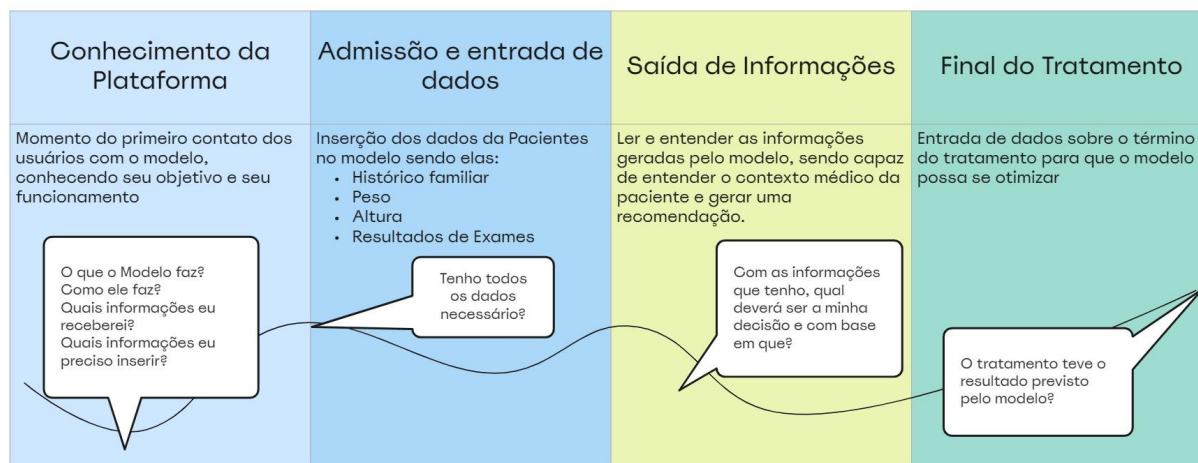


### Marcos Fernandes Neto - Médico

**Cenário:** Com o diagnóstico do paciente, ele precisa saber qual o melhor tratamento para a mesma

### Expectativas

Ele espera ser capaz de entender informações que sejam conclusivas para que ele possa indicar o tratamento mais apropriado



### Oportunidades

- Ser uma plataforma de fácil e rápido entendimento para ágil aprendizado pelos usuários
- Ter um relatório detalhado para fundamentar a lógica do modelo e facilitar o entendimento do médico sobre o por quê do resultado

### Responsabilidades

Garantir que as informações sejam claras, facilmente inseridas e entendidas pelo corpo médico

miro

Fonte: Elaboração própria



## 4.2. Compreensão dos Dados

### 1. Exploração de dados:

*Apresentar a estatística descritiva básica de cada coluna, identificação se a coluna é numérica ou categórica e pelo menos 3 gráficos para visualizar a relação entre variáveis escolhidas pelo grupo.*

- a) Cite quais são as colunas numéricas e categóricas.
- b) Estatística descritiva das colunas.

### 2. Pré-processamento dos dados:

*Realizar a limpeza (tratamento de missings e remoção de outliers) e transformação (normalização e codificação) das colunas.*

- a) Cite quais são os outliers e qual correção será aplicada.

### 3. Hipóteses:

*Três hipóteses sobre a relação dos dados e o problema.*

- a) Levantamento das três hipóteses com justificativa.

## 4.3. Preparação dos Dados e Modelagem

### Caso seu projeto seja:

#### 1. Modelo supervisionado:

- a) Modelagem para o problema (proposta de features com a explicação completa da linha de raciocínio).
- b) Métricas relacionadas ao modelo (conjunto de testes, pelo menos 3).
- c) Apresentar o primeiro modelo candidato, e uma discussão sobre os resultados deste modelo (discussão sobre as métricas para esse modelo candidato).

### Caso seu projeto seja:

#### 1. Modelo não-supervisionado:

- a) Modelagem para o problema (proposta de features com a explicação completa da linha de raciocínio).
- b) Primeiro modelo candidato para o problema.
- c) Justificativa para a definição do K do modelo.
- d) Escolha de um tipo de sistema de recomendação e a justificativa para essa escolha.

## 4.4. Comparação de Modelos

*- Escolha da métrica do modelo baseado no que é mais importante para o problema ao se medir a qualidade do modelo;*

*- Pelo menos três modelos candidatos com tuning de hiperparâmetros e suas respectivas métricas;*

*- Definição do modelo escolhido e justificativa.*

a) Escolha da métrica e justificativa.

b) Modelos otimizados.

- Apresentar três modelos e suas métricas.

- Os modelos apresentados foram otimizados utilizando algum algoritmo de otimização para os hiperparâmetros? Ex. Grid Search e Random Search.

c) Definição do modelo escolhido e justificativa.

## 4.5. Avaliação

*Descreva a solução final de modelo preditivo e justifique a escolha. Alinhe sua justificativa com a Seção 4.1, resgatando o entendimento do negócio e explicando de que formas seu modelo atende os requisitos. Descreva também um plano de contingência para os casos em que o modelo falhar em suas previsões.*

*Além disso, discuta sobre a explicabilidade do modelo e realize a verificação de aceitação ou refutação das hipóteses.*

*Se aplicável, utilize equações, tabelas e gráficos de visualização de dados para melhor ilustrar seus argumentos.*



## 5. Conclusões e Recomendações

*Escreva, de forma resumida, sobre os principais resultados do seu projeto e faça recomendações formais ao seu parceiro de negócios em relação ao uso desse modelo. Você pode aproveitar este espaço para comentar sobre possíveis materiais extras, como um manual de usuário mais detalhado na seção “Anexos”.*

*Não se esqueça também das pessoas que serão potencialmente afetadas pelas decisões do modelo preditivo e elabore recomendações que ajudem seu parceiro a tratá-las de maneira estratégica e ética.*

## 6. Referências

*Incluir as principais referências de seu projeto, para que seu parceiro possa consultar caso ele se interessar em aprofundar.*

*Um exemplo de referência de livro:*

*LUCK, Heloisa. **Liderança em gestão escolar**. 4. ed. Petrópolis: Vozes, 2010.*

*SOBRENOME, Nome. **Título do livro**: subtítulo do livro. Edição. Cidade de publicação: Nome da editora, Ano de publicação.*

## Anexos

*Utilize esta seção para anexar materiais como manuais de usuário, documentos complementares que ficaram grandes e não couberam no corpo do texto etc.*