

UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

LEONARDO TRES MARTINEZ

**Avaliação das taxas de cesariana ao longo dos anos em instituições públicas e privadas do município de Ribeirão Preto (SP)**

Ribeirão Preto – SP

2021

LEONARDO TRES MARTINEZ

**Avaliação das taxas de cesariana ao longo dos anos em instituições públicas e privadas do município de Ribeirão Preto (SP)**

Trabalho de Conclusão de Curso apresentado ao curso de Graduação em Informática Biomédica, como parte dos requisitos necessários à obtenção do título de Informata Biomédico.

Orientador: Lívia Maria de Oliveira Ciabati

Coorientador: Domingos Alves

Ribeirão Preto – SP

2021

FMRP – Faculdade de Medicina de Ribeirão Preto

Leonardo Tres Martinez

Esta Monografia foi julgada adequada para a obtenção do título de Bacharel em Informática Biomédica, sendo aprovada em sua forma final pela banca examinadora:

**Data da Avaliação:**

**29/07/2021**

---

Dra. Livia Maria de Oliveira Ciabati

---

Prof. Dr. Domingos Alves

---

Dra. Hayala Cristina Cavenague de Souza

Dedico a minha família, amigos e a Doutora  
Lívia Ciabati por todo o apoio que me foi dado  
durante o desenvolvimento deste trabalho.

## RESUMO

As taxas de cesárea continuam a aumentar tanto em países desenvolvidos quanto em países em desenvolvimento. O uso excessivo e desnecessário desse procedimento pode trazer diversos riscos à saúde da mulher. A Organização Mundial de Saúde (OMS) concluiu em 1985 que não havia motivo para qualquer região ter uma taxa de cesariana superior a 10-15%. No Brasil, todos os nascimentos devem ser registrados no Sistema de Informação de Nascidos Vivos (SINASC). Esse banco foi tratado e enriquecido pela Plataforma de Ciência de Dados Aplicada a Saúde (PCDaS). Essa plataforma também chama a atenção para as possibilidades da abordagem da Ciência de Dados para análise, monitoramento, predição de eventos e situações de saúde e doença na população. Nesse contexto, o presente estudo tem como objetivo a exploração de variações nas taxas de cesariana, a partir do banco anual proveniente do SINASC. Foram analisadas as taxas de cesárea em maternidades públicas e privadas do município de Ribeirão Preto. As taxas encontradas na cidade de 1996 a 2017 estão em média 57%. O excesso de cesarianas fica ainda mais evidente ao analisar apenas os hospitais particulares, em que até 2013, a cada 10 partos, apenas 1 era vaginal. Também foram analisadas as características demográficas e socioeconômicas das mulheres de acordo com o perfil da entidade de ocorrência do parto. As mulheres que tiveram partos no setor privado eram, em maior proporção, brancas, com maior nível de escolaridade e pertenciam as faixas etárias mais elevadas. Junto às análises também foram treinados modelo de aprendizado de máquina para a classificação de partos. Os modelos classificaram o perfil institucional (público ou privado) do hospital de acordo com as características das mulheres e o tipo de parto realizado (vaginal ou cesáreo). Os padrões de mulheres que tiveram partos em hospitais privados e públicos foram reconhecidos pelos modelos, que tiveram em sua maioria mais de 80% de acurácia. A partir dos resultados, é correto afirmar que a taxa de cesariana na cidade de Ribeirão Preto ainda está muito acima do nível recomendado pela OMS, principalmente em hospitais particulares. Portanto, é necessário implantar um sistema de monitoramento efetivo de cesáreas, principalmente em grupos de mulheres do setor privado de saúde.

Palavras-Chaves: Sistemas de Saúde; Cesárea; Ciência de dados; Aprendizado de máquina

## **ABSTRACT**

Caesarean rates continue to rise in both developed and developing countries. The excessive and unnecessary use of this procedure can bring several risks to women's health. The World Health Organization (WHO) concluded in 1985 that there was no reason for any region to have a caesarean rate higher than 10-15%. In Brazil, all births must be registered in the Live Birth Information System (SINASC). This database was treated and enriched by the Applied Health Data Science Platform (PCDaS). This platform also draws attention to the possibilities of the Data Science approach for analyzing, monitoring, predicting health and disease events in the population. In this context, this study aims to explore variations in cesarean rates, based on the annual data from SINASC. Cesarean rates in public and private maternity hospitals in the city of Ribeirão Preto were analyzed. The rates found in the city from 1996 to 2017 are 57% on average. The excess of cesarean sections is even more evident when analyzing only private hospitals, in which until 2013 of every 10 births, only 1 was vaginal. The demographic and socioeconomic characteristics of the women were also analyzed according to the profile of the entity where the birth occurred. Women who gave birth in the private sector were, in greater proportion, white, with a higher level of education and belonged to the older age groups. Along with the analyses, a machine learning model for classification of births was also trained. The models classified the institutional profile (public or private) of the hospital according to the characteristics of the women and the type of birth performed (vaginal or cesarean). The patterns of women who delivered in private and public hospitals were recognized by the models, which were mostly more than 80% accurate. Based on the results, it is correct to state that the cesarean rate in the city of Ribeirão Preto is still far above the level recommended by the WHO, especially in private hospitals. Therefore, it is necessary to implement an effective monitoring system for cesarean sections, especially in groups of women in the private health sector.

**Keywords:** Health Systems; Cesarean section; Data science; Machine learning

## Lista de figuras

<b>Figura 1</b> – Exemplo de Árvore de Decisão .....	15
<b>Figura 2</b> - Exemplo de classificação do kNN com dois rótulos de classe e $k = 7$ .....	16
<b>Figura 3</b> - Exemplos de hiperplanos separando os dados.....	18
<b>Figura 4</b> - Diagrama simplificado de uma Rede neural.....	19
<b>Figura 5</b> - Fluxograma do estudo.....	24

## **Lista de tabelas**

<b>Tabela 1</b> - Taxa de cesárea por ano em Ribeirão Preto .....	25
<b>Tabela 2</b> -Perfil e número de registros de partos das entidades escolhidas .....	26
<b>Tabela 3</b> - Distribuição do tipo de parto por ano e perfil da entidade.....	26
<b>Tabela 4</b> - Características das mulheres de acordo com o perfil da entidade de ocorrência do parto. Ribeirão Preto, 2006-2017.....	27
<b>Tabela 5</b> - Resultados dos algoritmos.....	28



## Sumário

<b>1. INTRODUÇÃO .....</b>	<b>9</b>
<b>2. OBJETIVOS .....</b>	<b>11</b>
2.1 OBJETIVO GERAL .....	11
2.2 OBJETIVOS ESPECÍFICOS .....	11
<b>3. REFERENCIAL TEÓRICO .....</b>	<b>12</b>
3.1 APRENDIZADO DE MÁQUINA.....	12
3.2 ALGORITMOS DE CLASSIFICAÇÃO.....	12
<b>4. MATERIAIS E MÉTODOS.....</b>	<b>20</b>
4.1. FONTE DE DADOS .....	20
4.2 VARIÁVEIS DE INTERESSE .....	20
4.3. MÉTODOS ESTATÍSTICOS .....	21
4.4. COMITÊ DE ÉTICA.....	22
<b>5. RESULTADOS .....</b>	<b>23</b>
<b>6. DISCUSSÃO .....</b>	<b>29</b>
<b>7. CONCLUSÃO.....</b>	<b>32</b>
<b>8. REFERÊNCIAS .....</b>	<b>33</b>

## 1. INTRODUÇÃO

As taxas de cesárea continuam a aumentar tanto em países desenvolvidos quanto em países em desenvolvimento. Quando indicadas por motivos médicos, as cesarianas podem efetivamente evitar a mortalidade materna e perinatal, porém o uso desnecessário e excessivo desse procedimento não traz os mesmos benefícios. A cirurgia pode trazer diversos riscos à saúde da mulher, que aumentam em populações com acesso limitado a cuidados obstétricos abrangentes (BETRAN et al., 2015, 2016)

No Brasil as taxas de cesarianas tem crescido consistentemente, mesmo muitas mulheres preferindo o parto vaginal à cesárea, principalmente as que já experimentaram as duas formas (FAÚNDES et al., 2004). Em 2009, pela primeira vez, o número de nascidos no país por meio cirúrgico foi maior do que pelo parto vaginal (DO CARMO LEAL et al., 2012). Estudos mostram que a escolha pela cesárea eletiva muitas vezes tem relação com o medo das violências obstétricas que podem ocorrer no parto vaginal. Experiências anteriores positivas com a cesárea também influenciaram a escolha de mães no setor privado, mas esse número de relatos foi significativamente menor no setor público (DOMINGUES et al., 2014).

O aumento injustificável da taxa de cesariana já vem sendo discutido há muito tempo. Em 1985 um grupo de experts foi convocado pela Organização Mundial de Saúde (OMS) para discutir sobre tecnologias apropriadas para o nascimento. Nesse evento, um dos assuntos discutidos foi: qual seria a taxa apropriada de cesárea em uma população. Visto que alguns dos países de menor índice de mortalidade perinatal na época possuíam uma taxa inferior a 10%, concluíram que não havia motivo para qualquer região ter uma superior a 10-15% (“Appropriate technology for birth”, 1985).

Porém, nos últimos anos, a OMS definiu o Sistema de Classificação de Robson (SCR) como um padrão para avaliar e comparar taxas de cesariana. Essa Classificação divide as mulheres em 10 grupos obstétricos totalmente inclusivos a partir de características obstétricas semelhantes, como a paridade (nulíparas, multíparas, com e sem cesárea anterior), o início de trabalho de parto (espontâneo, induzido ou cesárea anteparto), idade gestacional, apresentação fetal e número de fetos (ROBSON, 2001). A análise de cada um dos grupos pode ser uma alternativa para otimizar o uso da cesárea, focando intervenções em grupos específicos para cada unidade de saúde (BETRAN et al., 2016).

No Brasil, todos os nascimentos devem ser registrados no Sistema de Informação de Nascidos Vivos (SINASC). O DATASUS (“DATASUS”, [s.d.]) desenvolveu o SINASC visando reunir informações epidemiológicas referentes aos nascimentos informados em todo o país, contando hoje com um grande volume de dados provenientes de décadas de coleta. O banco foi implantado oficialmente a partir de 1990 e hoje seus dados têm grande importância para a gestão em saúde do país. Os dados podem ser usados para formulação de indicadores epidemiológicos como instrumentos estratégicos de suporte ao planejamento das ações, atividades e programas voltados à gestão em saúde.

A Plataforma de Ciência de dados Aplicada a Saúde (PCDaS) tratou e enriqueceu os dados obtidos do DATASUS, resultando no banco de dados usado para a análise nesse projeto. A PCDaS tem como objetivo principal disponibilizar serviços tecnológicos e computação científica para armazenamento, gestão e análise de grandes quantidades de dados (“Sistema de Informações de Nascidos Vivos - SINASC - Conjuntos de dados - PCDaS”, [s.d.]). Essa plataforma chama a atenção para as possibilidades da abordagem da Ciência de Dados para análise, monitoramento, predição de eventos e situações de saúde e doença na população. Hoje a PCDaS conta com 5 conjunto de dados: Sistema de Informações Hospitalares do Sistema Único de Saúde (SIHSUS), Cadastro Nacional de Estabelecimentos de Saúde (CNES), Sistema de Informações de Nascidos Vivos (SINASC), Sistema de Informações de Mortalidade (SIM) e Indicadores Programa Mais Médicos (PMM).

Dessa forma, o presente estudo tem como objetivo geral a exploração de variações nas taxas de cesariana, a partir do banco anual proveniente do SINASC. Além disso, foi avaliada a incidência de cesáreas de acordo com o perfil de pagamento (público ou privado) dos hospitais de Ribeirão Preto-SP.

## **2. OBJETIVOS**

### **2.1 OBJETIVO GERAL**

Analisar as diferenças nas taxas de cesariana nas maternidades públicas e privadas de Ribeirão Preto (SP) ao longo dos anos e reconhecer os padrões de perfil das mulheres nessas maternidades.

### **2.2 OBJETIVOS ESPECÍFICOS**

2.2.1 Analisar a população do banco de dados e identificar as instituições de saúde públicas e privadas do município de Ribeirão Preto;

2.2.2 Comparar a taxa de cesárea ao longo dos anos em que houve registros de partos nas maternidades escolhidas;

2.2.3 Treinar modelos de classificação usando algoritmos de aprendizado de máquina para identificar o perfil da maternidade em que o parto ocorreu, de acordo com as características da mãe e o tipo de parto.

2.2.4 Avaliar qual dos algoritmos foi o mais eficiente para a classificação e reconhecer os padrões encontrados.

### 3. REFERENCIAL TEÓRICO

#### 3.1 APRENDIZADO DE MÁQUINA

O Aprendizado de Máquina (AM) é um ramo de algoritmos computacionais projetados para emular a inteligência humana aprendendo com o ambiente a sua volta. Eles são a maior força de trabalho da nova era dos chamados *Big Data*. Técnicas baseadas em aprendizado de máquina têm sido aplicadas com sucesso em diversas áreas, desde reconhecimento de padrões, visão computacional, engenharia de naves espaciais, finanças, entretenimento e biologia computacional para aplicações biomédicas e médicas. (EL NAQA; MURPHY, 2015)

As técnicas de AM utilizam o princípio de indução, no qual obtém-se conclusões genéricas a partir de um conjunto de exemplos. O aprendizado indutivo pode ser dividido em dois tipos: supervisionado e não-supervisionado.

O aprendizado supervisionado também é referido como “aprender com um professor”. O professor fornece o conhecimento para o algoritmo, através de exemplos na forma de *input-output* (entrada/saída) (HAYKIN et al., 2009). As variáveis previsoras são as entradas do exemplo passado e o rótulo de classe é a saída do algoritmo. O algoritmo de AM extrai a representação do conhecimento a partir desses exemplos com o objetivo de que a representação gerada seja capaz de produzir saídas corretas para novas entradas (LORENA; DE CARVALHO, 2007).

Já no aprendizado não-supervisionado, também conhecido como “aprender sem um professor”, como o nome implica, não há um professor para supervisionar o processo de aprendizagem do algoritmo. Ou seja, não são passados exemplos rotulados para serem aprendidos (HAYKIN et al., 2009).

#### 3.2 ALGORITMOS DE CLASSIFICAÇÃO

O tipo de aprendizado abordado nesse trabalho é o supervisionado, mais em específico, a classificação. Na classificação, o objetivo de um algoritmo de AM é construir um classificador (ou modelo, preditor) dado um conjunto de exemplos de treinamento com seus rótulos de classes. Um classificador é uma função que atribui um rótulo de classe a um exemplo que possui variáveis (*features*) que o caracteriza. (ZHANG, 2004)

Neste trabalho, foram utilizados os seguintes algoritmos de classificação: *Naive Bayes*, *Árvore de Decisão*, *Random Forest*, *kNN*, *Regressão logística*, *Support Vector Machines* e *Redes neurais artificiais*. Os detalhes de cada uma dessas abordagens são apresentados nos tópicos a seguir.

### 3.2.1 *Naive Bayes*

O algoritmo de classificação *Naive Bayes* é baseado no “Teorema de Bayes”, o qual foi criado por Thomas Bayes (1701 – 1761). Seus métodos de aprendizado consistem em aplicar o teorema com a suposição “*naive*” (ingênua) de independência condicional entre as *features* dado o valor da variável de classe. O teorema de bayes afirma a seguinte relação, dada a variável de classe  $y$  e o vetor de características dependente  $x_1$  até  $x_n$ :

$$P(y|x_1, \dots, x_n) = \frac{P(y)P(x_1, \dots, x_n|y)}{P(x_1, \dots, x_n)}$$

Usando a suposição ingênua de independência condicional de que

$$P(x_i|y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = P(x_i|y),$$

Para todo  $i$ , essa relação é simplificada para

$$P(y|x_1, \dots, x_n) = \frac{P(y) \prod_{i=1}^n P(x_i|y)}{P(x_1, \dots, x_n)}$$

Já que  $P(x_1, \dots, x_n)$  é uma constante dada pelo exemplo, podemos usar a seguinte regra de classificação:

$$P(y|x_1, \dots, x_n) \propto P(y) \prod_{i=1}^n P(x_i|y)$$

↓

$$\hat{y} = \arg \max_y P(y) \prod_{i=1}^n P(x_i|y)$$

E podemos usar a estimação Máximo A Posteriori (MAP) para estimar  $P(y)$  e  $P(x_i|y)$ ; o primeiro é então a frequência relativa da classe  $y$  no conjunto de treinamento.

Os diferentes classificadores *Naive Bayes* se diferem principalmente pelas suposições que fazem sobre a distribuição de  $P(x_i|y)$ .

Este trabalho utiliza a implementação do algoritmo *Gaussian Naive Bayes*, com a função *GaussianNB* da biblioteca *scikit-learn*. Nesse algoritmo a probabilidade das *features* é assumida como gaussiana:

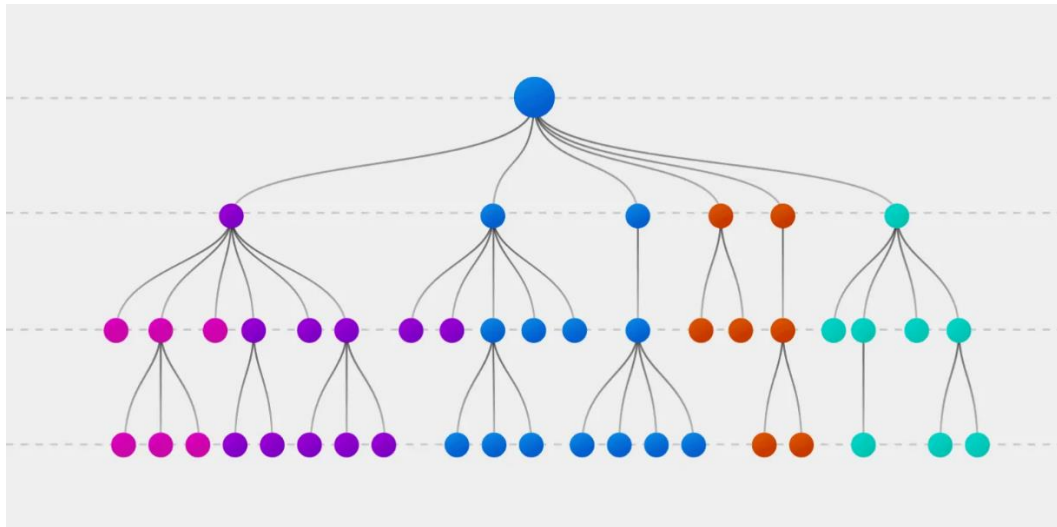
$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right)$$

Os parâmetros  $\sigma_y$  e  $\mu_y$  são estimados usando máxima verossimilhança (“1.9. Naive Bayes — scikit-learn 0.24.2 documentation”, [s.d.]).

### 3.2.2 Árvore de Decisão

O objetivo de uma Árvore de Decisão é criar um modelo que preveja o valor da variável de classe aprendendo regras de decisão simples inferidas a partir das *features* dos dados (“1.10. Decision Trees — scikit-learn 0.24.2 documentation”, [s.d.]).

O algoritmo foi implementado usando a função *DecisionTreeClassifier* da biblioteca *scikit-learn*. Essa função cria um modelo de Árvore de Decisão capaz de realizar a classificação multiclasse em um conjunto de dados. Nessa Árvore, os nós folhas contém a resposta (variável de classe) e os outros nós contém as variáveis previsoras (*features*). A Figura 1 mostra um exemplo de Árvore de Decisão.

**Figura 1** – Exemplo de Árvore de Decisão

Fonte: Imagem retirada do post "The Complete Guide to Decision Trees" do site [explorium.ai](http://explorium.ai)

### 3.2.3 *Random Forest*

As *Random Forests* são combinações de Árvore de Decisão, de modo que cada árvore depende dos valores de um vetor aleatório de *features* amostrado de forma independente (BREIMAN, 2001). Então, diferentemente da Árvore de Decisão, o algoritmo cria não uma, mas várias árvores que classificam uma classe cada uma.

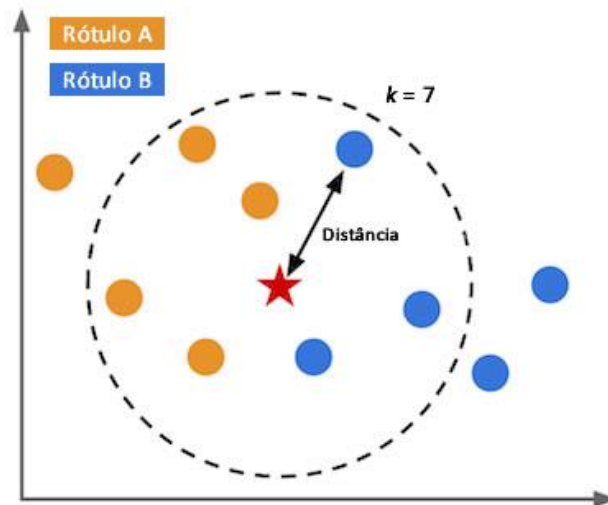
Neste trabalho foi utilizada a função *RandomForestClassifier* da biblioteca *scikit-learn*. Na implementação do *scikit-learn*, os classificadores são combinados pela média de suas previsões probabilísticas, em vez de deixar que cada classificador vote para uma única classe ("1.11. Ensemble methods — scikit-learn 0.24.2 documentation", [s.d.]).

### 3.2.4 kNN (K Vizinhos Mais Próximos)

O princípio do algoritmo kNN (ou K Vizinhos Mais Próximos) é encontrar um número de amostras de treinamento mais próximas do novo ponto e prever a classe a partir deles. O kNN utilizado neste trabalho está implementado na função *KNeighborsClassifier*, da biblioteca *scikit-learn*. O número k de amostras vizinhas pode ser determinado como um parâmetro da função. Na figura 2, pode-se ver um exemplo com  $k = 7$ , neste exemplo, a nova amostra (representada por uma estrela) receberia o rótulo A.



**Figura 2** - Exemplo de classificação do kNN com dois rótulos de classe e  $k = 7$



Fonte: Imagem retirada no post "K vizinhos mais próximos - KNN" do blog [computacaointeligente.com.br](http://computacaointeligente.com.br)

A classificação desse algoritmo é um tipo de aprendizado baseado em instâncias. Isso significa que, diferente dos algoritmos citados anteriormente, esse não constrói um modelo de predição, mas apenas armazena as instâncias dos dados de treinamento. A classificação é então definida a partir da maioria dos votos dos vizinhos mais próximos de cada ponto ("1.6. Nearest Neighbors — scikit-learn 0.24.2 documentation", [s.d.]).

### 3.2.5 Regressão logística

Regressão logística é um modelo linear para classificação, no qual as probabilidades que descrevem os resultados possíveis de um evento são modeladas usando uma função logística. Esse algoritmo fornece os resultados da análise em termos de probabilidade, então eles ficam contidos no intervalo de zero a um.

Podendo a variável de classe  $y$  assumir dois estados (1 ou 0, neste trabalho, público ou privado) e havendo um conjunto de  $n$  *features*  $X_1, X_2, X_3, \dots, X_n$ , o modelo pode ser escrito da seguinte forma:

$$P(Y = 1) = \frac{1}{1 + e^{-g(x)}}$$

Em que,

$$g(x) = B_0 + B_1X_1 + \dots + B_nX_n$$

Os coeficientes  $B_0, B_1, \dots, B_n$  são estimados a partir do conjunto de dados de treinamento pelo método de máxima verossimilhança. Dessa forma, o modelo acha uma combinação de coeficientes que maximiza a probabilidade de um exemplo ter uma determinada classe.

Portanto, quando  $g(x) \rightarrow +\infty$ , então  $P(Y = 1) \rightarrow 1$ , e quando  $g(x) \rightarrow -\infty$ , então  $P(Y = 1) \rightarrow 0$ .

Esse algoritmo também é implementado na biblioteca *scikit-learn*, usando a função *LogisticRegression* (“1.1. Linear Models — scikit-learn 0.24.2 documentation”, [s.d.]).

### 3.2.6 Support Vector Machine (SVM)

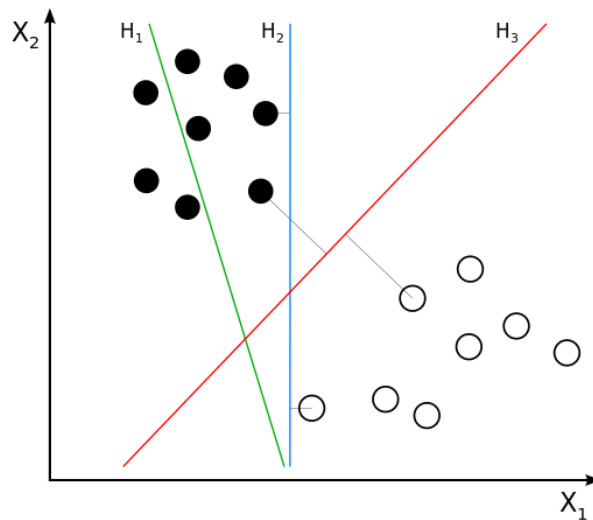
A ideia básica dos *Support Vector Machines* (SVMs) é encontrar um hiperplano que separa perfeitamente os dados em duas classes. Porém, como os dados nem sempre são linearmente separáveis, os SVMs precisam levar os dados em uma dimensão maior, na qual são separáveis (BOSWELL, 2002).

O algoritmo desta análise foi implementado usando a função SVC da biblioteca *scikit-learn*. Os métodos de classificação dos SVMs usam um subconjunto de pontos na função da decisão (chamados vetores de suporte). (“1.4. Support Vector Machines — scikit-learn 0.24.2 documentation”, [s.d.]). A distância entre esses pontos e o hiperplano separador é chamado de margem.

É possível ter mais de um hiperplano que separa as classes. Pode-se olhar para o que tem a maior distância entre as duas classes, ou seja, a maior margem, para escolher o melhor hiperplano. Então o melhor hiperplano seria aquele cujas distâncias entre ele e os primeiros vetores de suporte de cada classe sejam maximizadas.

Na figura 3, pode-se ver um exemplo de três hiperplanos (H1, H2 e H3) separando dados de duas classes diferentes (pontos pretos e pontos brancos). No exemplo, o hiperplano H1 não separa as duas classes, enquanto o H2 separa, porém com uma pequena margem. Então, o H3 seria o melhor hiperplano, pois separa os dados e suas diferentes classes com a maior margem.

**Figura 3** - Exemplos de hiperplanos separando os dados



Fonte: Imagem retirada da página "Support-vector machine" do site [en.wikipedia.org/wiki/Support-vector\\_machine](http://en.wikipedia.org/wiki/Support-vector_machine)

### 3.2.7 Redes neurais artificiais

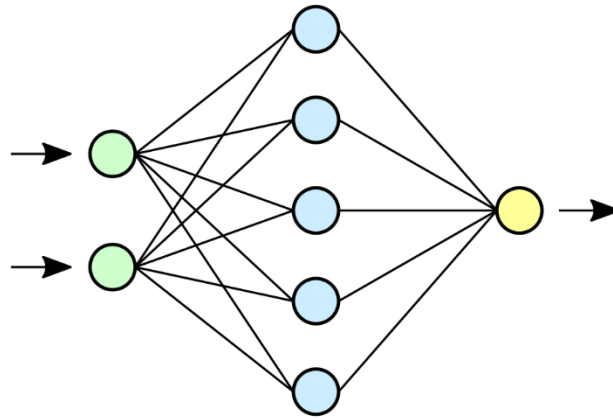
As Redes neurais artificiais têm como objetivo criar um sistema computacional para resolver problemas imitando um cérebro humano. O algoritmo que usa Redes neurais chamado de *Multi-layer Perceptron* (MLP) foi aplicado neste trabalho por meio da função *MLPClassifier* da biblioteca *scikit-learn*.

Uma Rede neural tem diversas unidades de processamento chamadas neurônios que podem ser divididos em três tipos de camadas: camada de entrada, camadas intermediárias e camada de saída. O diagrama simplificado de uma Rede neural apresentado na Figura 4 mostra essas camadas conectadas por linhas, representando os sinais que estão sendo passados de neurônio a neurônio. Cada uma das linhas de conexão entre as unidades tem um determinado peso, definido no processo de treinamento da Rede.

As Redes aprendem através dos exemplos dados, ajustando os pesos de suas conexões de acordo com os padrões apresentados. As *features* dos exemplos são passadas na camada de entrada e o valor da camada de saída é comparado ao rótulo de classe verdadeiro do exemplo. Os pesos das conexões entre as camadas intermediárias são então ajustados para que o erro da Rede (camada de saída ser diferente da classe verdadeira) seja minimizado.

Após o aprendizado, o modelo de Rede neural pode ser usado para classificar novos exemplos, utilizando os pesos ajustados no treinamento.

**Figura 4** - Diagrama simplificado de uma Rede neural



Fonte: Imagem retirada da página "Rede neural artificial" do site [pt.wikipedia.org/wiki/Rede\\_neural\\_artificial](http://pt.wikipedia.org/wiki/Rede_neural_artificial)

## 4. MATERIAIS E MÉTODOS

Este trabalho trata-se de um estudo descritivo-analítico cujo objetivo é avaliar a taxa de cesariana nas instituições públicas e privadas de Ribeirão Preto (SP).

### 4.1. FONTE DE DADOS

A informação a respeito dos eventos de gestação e parto foram obtidos a partir da PCDaS, que tratou e enriqueceu os dados do SINASC. Foram selecionadas as instituições de saúde da cidade de Ribeirão Preto, divididas entre públicas e privadas. Informações a respeito do perfil de cada uma das maternidades para essa classificação foram obtidas por meio do Cadastro Nacional de Estabelecimentos de Saúde (CNES). O critério para a definição dos perfis foi baseado na proporção de leitos do Sistema Único de Saúde (SUS) e leitos privados que a maternidade possui. Portanto, foram classificadas como públicas aquelas instituições com pelo menos 85% dos leitos obstétricos sendo do SUS e, como privadas, aquelas com 100% de leitos obstétricos privados. Os dados da proporção de leitos foram obtidos no site do Ministério da saúde (MINISTÉRIO DA SAÚDE, 2015). Instituições que não se encaixavam em nenhum dos dois perfis foram definidas como mistas e não fizeram parte das análises finais. Da mesma forma, dados de partos não hospitalares foram excluídos, por não apresentarem risco de cesariana.

Os dados foram baixados pela primeira vez a partir da PCDaS no dia 8 de abril de 2021, com um total de 62.856.188 partos registrados de 1996 até 2017. Para a análise, o critério de inclusão foi o parto ter acontecido em uma das entidades escolhidas na cidade de Ribeirão Preto. O viés do banco se dá principalmente a erros de anotação, com algumas variáveis sendo anotadas erroneamente. Porém, valores inválidos já foram eliminados pelo tratamento da PCDaS.

### 4.2 VARIÁVEIS DE INTERESSE

As variáveis de desfecho para este estudo são: tipo de parto (vaginal ou cesáreo), presente originalmente na base coletada e o perfil da instituição de saúde (pública ou privada).

Já as variáveis independentes incluem: número de fetos (único ou múltiplos), paridade (primíparas ou múltiparas), cesárea anterior (com e sem cesariana prévia) e início do trabalho de parto (espontâneo, induzido ou cesariana eletiva). Além disso,

também foram obtidas variáveis que representam características demográficas e socioeconômicas da mãe, como: idade, escolaridade, etnia/cor da pele, estado civil e número de consultas pré-natal.

Para o treinamento dos modelos de classificação, foram utilizadas as seguintes variáveis referentes a mãe como variáveis previsoras: idade, etnia/cor da pele, estado civil, escolaridade e tipo de parto. Já para a variável de classe, foi utilizada a variável referente ao perfil da maternidade (público ou privado) em que o parto ocorreu.

#### 4.3. MÉTODOS ESTATÍSTICOS

Primeiramente, foi feita a análise descritiva da taxa de cesárea na cidade de Ribeirão Preto em cada ano do estudo. A seguir, foi feita a seleção das maternidades, a definição de seus perfis institucionais (público ou privado) e a contagem de registros de partos em cada uma. Então foi realizada a análise da taxa de cesárea por ano, de acordo com o perfil de cada entidade.

Foram descritas também as distribuições das mulheres pelas características demográficas e socioeconômicas de acordo com o perfil institucional da maternidade.

Por último, foram criados modelos de aprendizado de máquina para a classificação do perfil de entidade, de acordo com as características da mãe e o tipo de parto. Esse processo se dividiu em 4 etapas:

Coleta dos dados: foram coletados todos os dados do banco relativos as maternidades escolhidas no trabalho.

Pré-processamento dos dados: primeiramente, foi feita a escolha das variáveis previsoras e a criação de uma nova variável referente ao perfil institucional de cada maternidade. A nova variável foi definida como a variável de classe. Após isso, foi feito o tratamento de valores inconsistentes e valores faltantes das variáveis previsoras. Esse tratamento incluiu excluir os registros que tinham valores referentes a dados ignorados e retirar outliers. Foi feita a utilização da função *OneHotEncoder* para criar colunas binárias para cada valor de cada variável categórica utilizada. Também foi realizado o escalonamento dos valores de cada variável para normalizar as variáveis contínuas. Por fim, foi feita a divisão de bases em treinamento e teste, com 20% do total de dados para teste.

Treinamento e teste dos modelos: as bases de treinamento já pré-processadas foram passadas como parâmetro para as funções de cada algoritmo. Já com os modelos treinados, estes foram apresentados para a base de teste (contendo exemplos desconhecidos do modelo treinado) para que os modelos fizessem a previsão da variável de classe.

Avaliação dos algoritmos: os algoritmos foram avaliados pelas taxas de erro, sensibilidades e especificidades.

As análises dos dados e a construção e avaliação dos modelos foram realizadas utilizando a linguagem *python*, a *IDE Spyder* e a *IDE Jupyter Notebook*. Os dados de análise foram carregados em uma planilha do Excel para montar as tabelas e gráficos. O código utilizado nessa análise está publicado no *Github* (<https://github.com/leotmartinez/tcc>).

#### 4.4. COMITÊ DE ÉTICA

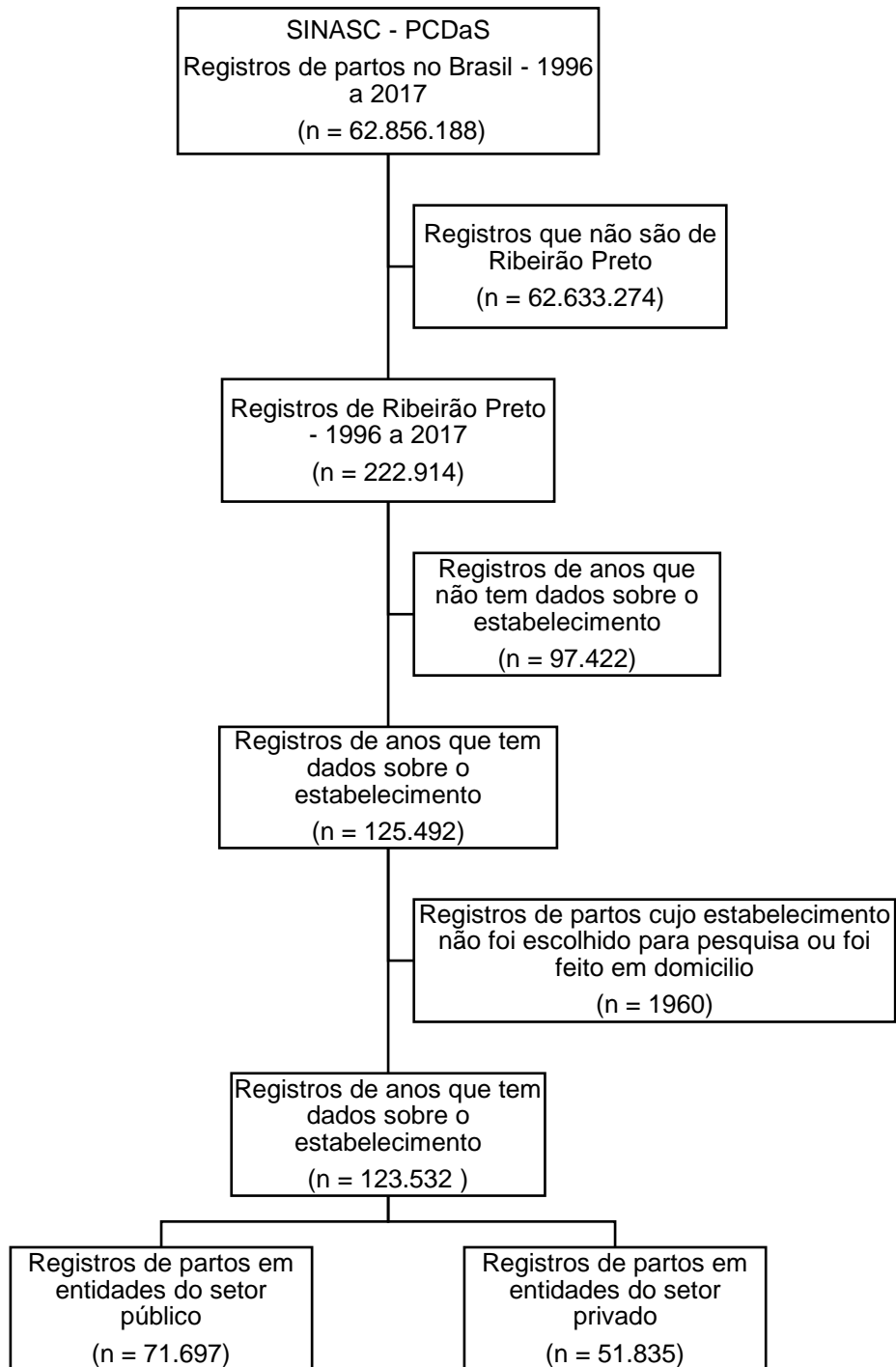
Como esse trabalho utiliza dados abertos, de acesso irrestrito e disponibilizados pelas próprias instituições, não foi necessária a aprovação pelo comitê de ética em pesquisa (CONSELHO NACIONAL DE SAÚDE, 2016).

## 5. RESULTADOS

A partir do banco de dados público (SINASC) disponibilizado pela PCDaS, foi feito o download dos dados de registros de partos no Brasil, de 1996 a 2017. Porém, foram excluídos da pesquisa os dados de 2015, pois nesse ano não foram encontrados dados suficientes da cidade de Ribeirão Preto para a análise. Dados referentes ao estabelecimento onde o parto ocorreu foram encontrados apenas nos bancos de 2006 a 2017.

Então, os dados de 62.856.188 registros de partos, oriundos de todas as cidades do Brasil entre 1996 e 2017 (menos 2015), foram incluídos na pesquisa. Desses, apenas 222.914 (0,35%;  $222.914/62.856.188$ ) foram registrados na cidade de Ribeirão Preto, dos quais 125.492 tinham dados referentes ao estabelecimento. Um total de 123.532 partos foram classificados como “público” ou “privado”, segundo o perfil do estabelecimento, sendo 71.697 públicos (0,11%;  $71.697/62.856.188$ ) e 51.835 privados (0,08%;  $51.835/62.856.188$ ). A figura 1 apresenta o fluxograma do estudo.



**Figura 5** - Fluxograma do estudo

Fonte: Autor 2021

A Tabela 1 apresenta a taxa de cesárea por ano em Ribeirão Preto. Do total de 222.917 registros de partos, 57% foram classificados como cesárea e 43% como vaginal. Em todos os anos, a taxa de partos cesáreos foi maior do que a taxa de partos vaginais. Começando em 56% no ano de 1996, a taxa de cesárea subiu e teve seu

maior valor em 2012, com 62% do total de partos sendo feitos por cesariana. Após 2012, essa taxa diminui nos anos seguintes, para seu último valor do conjunto de dados: 57% em 2017.

**Tabela 1 - Taxa de cesárea por ano em Ribeirão Preto**

<b>Ano</b>	<b>Total de partos</b>	<b>Vaginal</b>	<b>% Vaginal</b>	<b>Cesáreo</b>	<b>% Cesáreo</b>
1996	8.213	3.608	44%	4.605	56%
1997	9.846	4.296	44%	5.550	56%
1998	10.002	4.550	45%	5.452	55%
1999	10.068	4.667	46%	5.401	54%
2000	10.150	4.491	44%	5.659	56%
2001	9.628	4.095	43%	5.533	57%
2002	9.766	4.272	44%	5.494	56%
2003	9.720	4.277	44%	5.443	56%
2004	9.773	4.199	43%	5.574	57%
2005	10.256	4.407	43%	5.849	57%
2006	10.025	4.306	43%	5.719	57%
2007	9.974	4.106	41%	5.868	59%
2008	11.842	5.558	47%	6.284	53%
2009	10.747	4.206	39%	6.541	61%
2010	11.309	4.500	40%	6.809	60%
2011	11.787	4.659	40%	7.128	60%
2012	11.745	4.495	38%	7.250	62%
2013	11.644	4.569	39%	7.075	61%
2014	12.093	5.057	42%	7.036	58%
2016	11.956	5.558	46%	6.398	54%
2017	12.370	5.699	46%	6.671	54%
Total	222.914	95.575	43%	127.339	57%

A Tabela 2 apresenta todos os nomes das entidades escolhidas para a pesquisa, assim como seu perfil institucional (público, privado ou misto), os anos em que tiveram registros e o total de registros. Foram escolhidas 3 entidades privadas, com seus totais somados em 51.835 registros de partos e 5 entidades públicas, com 71.697 registros. Nenhuma das instituições foi definida como mista, portanto, todas participaram das análises.

**Tabela 2** -Perfil e número de registros de partos das entidades escolhidas

Nome	Perfil	Anos com registro	Total de registros
São Francisco Rede De Saúde Assistencial Ltda	Privada	2006 - 2017	45.034
Centro De Ref. Da Saúde Da Mulher De Ribeirão Preto - Mater	Pública	2006 - 2017	31.978
Hospital das Clínicas de Ribeirão Preto	Pública	2006 - 2017	20.290
Hospital Santa Casa de Misericórdia de Ribeirão Preto	Pública	2006 - 2017	14.817
Hospital São Paulo	Privada	2006 - 2017	3.947
Hospital Electro Bonini - Unaerp	Pública	2014 - 2017	2.952
Hospital São Lucas	Privada	2006 - 2017	2.854
Hospital Santa Lydia	Pública	2006 - 2017	1.660

A Tabela 3 apresenta a distribuição do tipo de parto por ano e perfil da entidade. No setor público, em todos os anos o número de cesarianas foi menor do que o de partos vaginais. Já no setor privado, ocorre o oposto, em todos os anos a taxa de cesariana foi maior que a taxa de partos vaginais, chegando até 91%. Mesmo o número total de partos no setor público sendo maior do que no privado (71.678 a 51.825), o número absoluto de cesarianas no setor privado é sempre maior do que no público. De 2014 para 2016 vemos a maior diminuição dessa taxa nos hospitais particulares, de 86% para 79%.

**Tabela 3** - Distribuição do tipo de parto por ano e perfil da entidade

Ano	Público (SUS)					Privado				
	Total de partos		Vaginal		Cesariana	Total de partos		Vaginal		Cesariana
	n		n	(%)	n (%)	n		n	(%)	n (%)
2006	6.024		3.920	65%	2.104 35%	3.971		361	9%	3.610 91%
2007	5.993		3.731	62%	2.262 38%	3.943		339	9%	3.604 91%
2008	6.125		3.731	61%	2.394 39%	4.334		446	10%	3.888 90%
2009	6.225		3.713	60%	2.512 40%	4.468		441	10%	4.027 90%
2010	6.523		4.021	62%	2.502 38%	4.747		441	9%	4.306 91%
2011	6.587		4.154	63%	2.433 37%	5.157		464	9%	4.693 91%
2012	6.644		3.950	59%	2.694 41%	5.050		496	10%	4.554 90%
2013	6.451		3.945	61%	2.506 39%	5.122		576	11%	4.546 89%
2014	6.726		4.236	63%	2.490 37%	5.246		747	14%	4.499 86%
2016	7.038		4.476	64%	2.562 36%	4.856		1.039	21%	3.817 79%
2017	7.342		4.580	62%	2.762 38%	4.931		1.043	21%	3.888 79%
Total	71.678		44.457	62%	27.221 38%	51.825		6.393	12%	45.432 88%

A tabela 4 apresenta as características das mulheres de acordo com o perfil institucional do hospital em que tiveram o parto, entre 2006 e 2017.

**Tabela 4** - Características das mulheres de acordo com o perfil da entidade de ocorrência do parto.  
Ribeirão Preto, 2006-2017.

Variáveis	Total		Público		Privado	
	n	%	n	%	n	%
<b>Idade (anos)</b>						
<20	15.845	12,8%	13.745	19,2%	2.100	4,1%
20-30	67.216	54,4%	41.027	57,2%	26.189	50,5%
>30	40.471	32,8%	16.925	23,6%	23.546	45,4%
<b>Cor da pele</b>						
Amarela	268	0,2%	92	0,1%	176	0,3%
Branca	92.873	75,4%	47.424	66,4%	45.449	87,8%
Indígena	12	0,0%	12	0,0%	0	0,0%
Parda	23.482	19,1%	18.673	26,1%	4.809	9,3%
Preta	6.544	5,3%	5.242	7,3%	1.302	2,5%
<b>Estado Civil</b>						
Casado	57.920	47,4%	21.624	30,6%	36.296	70,2%
União estável	11.970	9,8%	11.600	16,4%	370	0,7%
Separado Judic. /Divorciado	2.323	1,9%	1.399	2,0%	924	1,8%
Solteiro	49.806	40,7%	35.788	50,7%	14.018	27,1%
Viúvo	260	0,2%	160	0,2%	100	0,2%
<b>Escolaridade (anos)</b>						
<8	25.119	20,5%	22.735	32,0%	2.384	4,7%
8-11	64.957	53,1%	42.288	59,5%	22.669	44,2%
≥12	32.288	26,4%	6.085	8,6%	26.203	51,1%
<b>Paridade</b>						
Primípara	15.887	44,0%	8.171	38,7%	7.716	51,3%
Múltipara	20.252	56,0%	12.935	61,3%	7.317	48,7%
<b>Gestação Múltipla</b>						
Sim	3.912	3,2%	2.100	2,9%	1.812	3,5%
Não	119.617	96,8%	69.596	97,1%	50.021	96,5%
<b>Cesárea eletiva</b>						
Sim	20.811	59,3%	6.332	50,0%	14.479	64,6%
Não	14.254	40,7%	6.327	50,0%	7.927	35,4%
<b>Indução</b>						
Sim	9.852	16,1%	7.993	23,2%	1.859	6,9%
Não	51.358	83,9%	26.423	76,8%	24.935	93,1%
<b>Cesárea prévia</b>						
Sim	17.100	27,1%	8.517	23,5%	8.583	31,9%
Não	46.047	72,9%	27.697	76,5%	18.350	68,1%
<b>Pré-natal (n° consultas)</b>						
≤6	22.202	19,1%	17.690	25,3%	4.512	9,7%
≥7	94.295	80,9%	52.154	74,7%	42.141	90,3%
<b>Tipo de parto</b>						
Cesariana	72.653	58,8%	27.221	38,0%	45.432	87,7%
Vaginal	50.850	41,2%	44.457	62,0%	6.393	12,3%

Porcentagem de dados que não puderam ser obtidos para as seguintes variáveis: idade 0%, cor de pele 0,28%, estado civil 1%, escolaridade 0,9%, paridade 70,74%, gestação múltipla 0%, cesárea eletiva 51,52%, indução 50,45%, cesárea prévia 48,88%, pré-natal 5,69%, tipo de parto 0,02%.

Do total de 123.532 partos ocorridos em maternidades classificadas em públicas ou privadas, a taxa global de cesariana foi de 58,8% (Tabela 4). Porém, esse valor mostrou-se mais elevado entre as mulheres tendo parto no sistema privado, comparadas às do sistema público, 87,7% e 38% respectivamente.

As mulheres que tiveram parto no setor privado apresentaram proporções mais elevadas nos grupos de maior idade, com maior nível de escolaridade, maior proporção da pele branca, casadas e apresentaram uma maior frequência ao pré-natal. O setor público apresentou uma maior proporção de múltiparas (61,3% vs 48,7%) e uma proporção menor de mulheres com cesariana prévia (23,5% vs 31,9%). O uso da indução foi três vezes mais comum nas mulheres do setor público (23,2% vs 6,9%), enquanto as cesarianas eletivas foram mais frequentes no setor privado (64,6% vs 50%). Não foram encontradas grandes diferenças na proporção de partos gemelares entre os dois setores classificados pela forma de pagamento.

A respeito dos modelos de classificação, a tabela 5 mostra os resultados obtidos de todos os modelos usados. A acurácia de cada modelo mostra o quão bom ele foi em classificar qual é o perfil da entidade em que o parto ocorreu. Inversamente, a taxa de erro mostra o quão frequente o modelo erra sua classificação. Os valores de especificidade e sensibilidade mostram o quão certo o classificador está quando diz que o parto ocorreu em setor privado e público, respectivamente.

**Tabela 5** - Resultados dos algoritmos

	Acurácia (%)	Taxa de erro (%)	Sensibilidade (%)	Especificidade (%)
<i>Naive Bayes</i>	63,1	36,9	53	95
Árvore de Decisão	81,2	18,8	75	86
<i>Random Forest</i>	81,3	18,7	76	86
kNN	80,4	19,6	74	86
Regressão Logística	81,6	18,4	77	85
SVM	81,4	18,6	76	86
Redes neurais	81,7	18,3	76	87

## 6. DISCUSSÃO

Os achados desse estudo indicam que a taxa de cesariana na cidade de Ribeirão Preto ainda está muito acima do nível recomendado pela OMS, principalmente em hospitais particulares. Já em 1996, a maioria (56%) dos partos na cidade foram realizados por cesárea. Essa taxa foi mantida por décadas e, em todos os anos do estudo, o número de cesáreas foi maior do que o número de partos vaginais. O excesso fica ainda mais evidente ao analisar apenas os hospitais particulares, em que até 2013, a cada 10 partos, apenas 1 era vaginal.

As taxas de cesariana em Ribeirão Preto encontradas nos anos desse estudo estão em média 57%. Essa taxa é superior a taxa de cesariana do Brasil (55,5%), referida internacionalmente como uma das maiores do mundo, abaixo apenas da República Dominicana (58,1%) e de Chipre (56,9%) (WHO, 2008).

A respeito das características demográficas e socioeconômicas das mulheres, as que tiveram partos no setor privado eram, em maior proporção, brancas, com maior nível de escolaridade e pertenciam as faixas etárias mais elevadas. Essas características já foram associadas a uma maior taxa de cesariana em outros estudos no Brasil (RATTNER; MOURA, 2016). Este, porém, não é a única evidência que correlaciona as características hospitalares e obstétricas sobre a chance de realização de cesariana eletiva na Região Sudeste do Brasil. A chance de realizar uma cesariana eletiva é maior em um hospital particular, sendo 83,2% de prevalência de cesárea eletiva em hospitais privados (ZAIDEN et al., 2019). As análises desse trabalho corroboram com os resultados dos estudos referenciados. Os resultados mostraram que 64,6% das mulheres que tiveram parto em hospitais privados não entraram em trabalho de parto. Havia também mais mulheres multíparas e com um menor número de cesárea anterior no setor público. Com isso pode-se entender que a primeira cesárea tem uma influência em cesáreas posteriores. Assim como em outros estudos, pode-se concluir que as políticas públicas para diminuição de cesarianas eletivas devem ser dirigidas a mulheres primíparas. Em vista disso, ocorreria uma diminuição na cesárea de repetição em mulheres multíparas (NAKAMURA-PEREIRA et al., 2016).

Outra característica marcante das análises que predominou em mulheres de hospitais privados foi o maior número de consultas pré-natais. O Programa de

Humanização no Pré-natal e Nascimento (PHPN), do Ministério da Saúde, recomenda que devam ser realizadas, no mínimo, seis consultas de pré-natal (SERRUYA; CECATTI; LAGO, 2004). Esse resultado mostra um aparente paradoxo, pois com mais consultas, seria esperado que o parto normal fosse mais incentivado.

Um resultado importante de ser discutido foi a redução da taxa de cesariana em maternidades privadas entre os anos de 2014 e 2016. Essa redução pode ser explicada pela ocorrência de iniciativas para reduzir o número excessivo de cesáreas nesse período. Um bom exemplo foi o projeto Parto Adequado, que ocorreu entre 2015 e 2016 e contou com a participação de maternidades de todo o Brasil. Nesse projeto foram promovidas campanhas destinadas as mulheres, levando informações do não agendamento do parto e as vantagens do parto vaginal (NEGRÃO, 2017). Os resultados da primeira fase do projeto já mostraram grandes indicadores de diminuição no número de cesáreas desnecessárias e aumento no número de partos vaginais. Ao final de 18 meses, mais de 10 mil cesáreas sem indicação clínica foram evitadas nos hospitais participantes do projeto, mostrando sua efetividade.

Os resultados obtidos a partir do treinamento dos modelos de classificação, mostraram que, mesmo para o algoritmo, foi possível encontrar um padrão a respeito das variáveis previsoras e de classe. A maior parte dos modelos conseguiu prever mais de 80% dos perfis institucionais dos hospitais em que os partos ocorreram, apenas com a idade da mulher, seus anos de estudo, estado civil, cor da pele e o tipo de parto realizado. Pode-se notar que o modelo com maior taxa de erro, *Naive Bayes*, também teve a maior especificidade, porém menor sensibilidade entre os modelos. O algoritmo de *Naive Bayes* é conhecido por ter uma boa acurácia mesmo com poucos dados de treinamento, entretanto, ele desconsidera correlação entre as *features*, o que pode ter causado mais erros (ZHANG, 2004). O modelo de Redes neurais, por outro lado, foi o que apresentou os melhores resultados de classificação.

Um dos pontos fortes desse trabalho foi o uso de um grande banco de dados (SINASC) para as análises, além da utilização de um *script* automatizado de coleta dos dados. Desse modo foi possível reunir os dados de milhões de partos no Brasil e selecionar apenas os que seriam de interesse para a análise. Pela quantidade de dados analisados, seria inviável realizar uma coleta manual. Outro ponto forte está na origem dos dados, disponibilizados por uma fonte confiável (DATASUS) e que podem ser considerados de alta qualidade após serem tratados pela PCDaS. Da mesma

forma, o banco utilizado para a classificação dos perfis de cada entidade (CNES) é um sistema de informação do Ministério da Saúde, também podendo ser considerado de alta confiabilidade.

Entre os pontos fracos está a não utilização do Sistema de Classificação de Robson como principal parâmetro de avaliação da taxa de cesariana. Decerto com a utilização do SCR, poderia ser feita uma análise mais detalhada, focada na taxa de cesárea em cada um dos 10 grupos. Outro ponto negativo está na falta de dados de Ribeirão Preto no ano de 2015, além da falta de dados sobre o estabelecimento em todos os anos antes de 2006. Entretanto, mesmo com dados faltantes, a quantidade coletada para as análises ainda foi muito grande.

A acurácia dos modelos junto com os resultados das análises nas maternidades mostra que existe um padrão de mulheres que tem partos em hospitais do setor privado de saúde. Os resultados também mostram que esse padrão de mulheres são as que mais realizam cesarianas entre os dois setores. Portanto é importante que as iniciativas para a diminuição do parto cesáreo e a conscientização sobre o parto vaginal sejam direcionadas para esse grupo de mulheres.

Baseado no que foi discutido, os resultados desse trabalho podem impactar em outras pesquisas relacionadas às taxas de cesariana em maternidades de Ribeirão Preto. Compreendendo a epidemia de cesáreas na cidade e no país, podem ser descobertas novas opções de intervenções e consequentemente possibilitar o desenvolvimento de novas iniciativas para solucionar o problema. Também podem ser realizadas análises mais profundas em cada uma das maternidades escolhidas para esse projeto. Desse modo, seria possível entender em quais maternidades privadas esse problema é maior e desenvolver estratégias específicas para aquele local.



## 7. CONCLUSÃO

Certamente, bancos de dados como o SINASC apresentam um grande potencial informativo. Com a produção de dados crescendo ainda mais, é importante desenvolver estratégias de organização e contextualização de dados, para garantir a extração de informações relevantes para a solução de problemas. Os resultados desse trabalho buscam contribuir para o melhor entendimento da questão das altas taxas de cesariana, principalmente em hospitais do setor privado.

As altas taxas de cesariana encontradas são preocupantes, devido aos potenciais efeitos adversos para a mãe e o recém-nascido. Chama a atenção as taxas mais elevadas entre as mulheres do sistema privado, um resultado incoerente, pois são aquelas com um perfil socioeconômico privilegiado, no qual poderiam ser esperadas menores taxas.

Portanto, é necessário implantar um sistema de monitoramento efetivo de cesáreas, reduzindo as taxas principalmente em grupos de menor risco, como proposto pelo Sistema de Classificação de Robson (ROBSON, 2001). Também é importante apontar para iniciativas que visam diminuir o número excessivo de cesáreas no Brasil, como o projeto Parto Adequado.

## 8. REFERÊNCIAS

**1.1. Linear Models — scikit-learn 0.24.2 documentation.** Disponível em: <[https://scikit-learn.org/stable/modules/linear\\_model.html#logistic-regression](https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression)>. Acesso em: 23 jun. 2021.

**1.10. Decision Trees — scikit-learn 0.24.2 documentation.** Disponível em: <<https://scikit-learn.org/stable/modules/tree.html#tree>>. Acesso em: 23 jun. 2021.

**1.11. Ensemble methods — scikit-learn 0.24.2 documentation.** Disponível em: <<https://scikit-learn.org/stable/modules/ensemble.html>>. Acesso em: 23 jun. 2021.

**1.4. Support Vector Machines — scikit-learn 0.24.2 documentation.** Disponível em: <<https://scikit-learn.org/stable/modules/svm.html#svm-classification>>. Acesso em: 23 jun. 2021.

**1.6. Nearest Neighbors — scikit-learn 0.24.2 documentation.** Disponível em: <<https://scikit-learn.org/stable/modules/neighbors.html#id5>>. Acesso em: 23 jun. 2021.

**1.9. Naive Bayes — scikit-learn 0.24.2 documentation.** Disponível em: <[https://scikit-learn.org/stable/modules/naive\\_bayes.html#naive-bayes](https://scikit-learn.org/stable/modules/naive_bayes.html#naive-bayes)>. Acesso em: 23 jun. 2021.

Appropriate technology for birth. **Lancet**, v. 2, n. 8452, p. 436–437, 24 ago. 1985.

BETRAN, A. P. et al. What is the optimal rate of caesarean section at population level? A systematic review of ecologic studies. **Reproductive Health**, v. 12, n. 1, 21 jun. 2015.

BETRAN, A. P. et al. **WHO statement on caesarean section rates**BJOG: An International Journal of Obstetrics and GynaecologyBlackwell Publishing Ltd, , 1 abr. 2016.

BOSWELL, D. **Introduction to Support Vector Machines**. [s.l: s.n.].

BREIMAN, L. **Random Forests**. [s.l: s.n.].

CONSELHO NACIONAL DE SAÚDE. **Resolução nº 510/2016**Brasil, 2016.

**DATASUS.** Disponível em: <<http://www2.datasus.gov.br/DATASUS/index.php?area=060702>>. Acesso em: 11 maio. 2021.

DO CARMO LEAL, M. et al. **Birth in Brazil: National survey into labour and birth**Reproductive Health, 2012.

DOMINGUES, R. M. S. M. et al. **Processo de decisão pelo tipo de parto no Brasil: Da preferência inicial das mulheres à via de parto final**Cadernos de Saude Publica. [s.l: s.n.].

EL NAQA, I.; MURPHY, M. J. What Is Machine Learning? In: **Machine Learning in**

**Radiation Oncology.** [s.l.] Springer International Publishing, 2015. p. 3–11.

FAÚNDES, A. A. et al. Opinião de mulheres e médicos brasileiros sobre a preferência pela via de parto. **Revista de Saúde Pública**, v. 38, n. 4, p. 488–494, 2004.

HAYKIN, S. et al. **Neural Networks and Learning Machines Third Edition.** [s.l.: s.n.].

LORENA, A. C.; DE CARVALHO, A. C. P. L. F. Uma Introdução às Support Vector Machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43–67, 20 dez. 2007.

MINISTÉRIO DA SAÚDE. **Cadastro Nacional de Estabelecimentos de Saúde.** Disponível em: <<http://cnes.datasus.gov.br/pages/estabelecimentos/consulta.jsp>>. Acesso em: 15 maio. 2021.

NAKAMURA-PEREIRA, M. et al. Use of Robson classification to assess cesarean section rate in Brazil: The role of source of payment for childbirth. **Reproductive Health**, v. 13, n. Suppl 3, 17 out. 2016.

NEGRÃO, A. Iniciativas para diminuir o número de cesáreas excessivas no Brasil: Projeto Parto Adequado. 2017.

RATTNER, D.; MOURA, E. C. DE. Nascimentos no Brasil: associação do tipo de parto com variáveis temporais e sociodemográficas. **Revista Brasileira de Saúde Materno Infantil**, v. 16, n. 1, p. 39–47, 1 jan. 2016.

ROBSON, M. S. Can we reduce the caesarean section rate? **Best Practice & Research Clinical Obstetrics & Gynaecology**, v. 15, n. 1, p. 179–194, 1 fev. 2001.

SERRUYA, S. J.; CECATTI, J. G.; LAGO, T. DI G. DO. O Programa de Humanização no Pré-natal e Nascimento do Ministério da Saúde no Brasil: resultados iniciais. **Cadernos de Saúde Pública**, v. 20, n. 5, p. 1281–1289, 2004.

**Sistema de Informações de Nascidos Vivos - SINASC - Conjuntos de dados - PCDaS.** Disponível em: <<https://bigdata-metadados.iciict.fiocruz.br/dataset/sistema-de-informacoes-de-nascidos-vivos-sinasc>>. Acesso em: 17 maio. 2021.

WHO. **Births by Caesarean Section.** [s.l.: s.n.]. Disponível em: <[https://www.who.int/data/gho/data/indicators/indicator-details/GHO/births-by-caesarean-section-\(-\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/births-by-caesarean-section-(-))>. Acesso em: 12 jul. 2021.

ZAIDEN, L. et al. Influência das características hospitalares na realização de cesárea eletiva na Região Sudeste do Brasil. **Cadernos de saúde publica**, v. 36, n. 1, p. e00218218, 20 dez. 2019.

ZHANG, H. The Optimality of Naive Bayes Naive Bayes and Augmented Naive Bayes. **Aa**, v. 1, n. 2, p. 3, 2004.