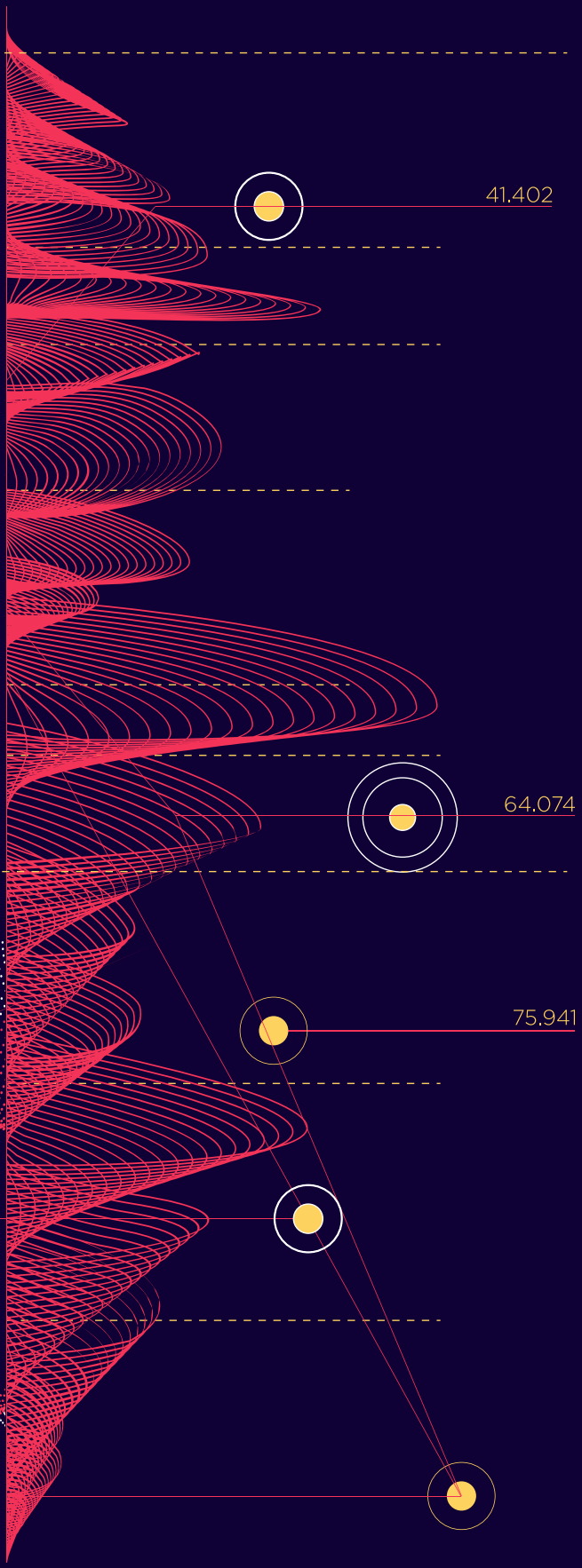
Relatório Metodológico

# MODELO DE CLASSIFICAÇÃO DE TECNOLOGIAS DE ACESSO

ceptro.br nic.br cgi.br

Relatório Metodológico

# MODELO DE CLASSIFICAÇÃO DE TECNOLOGIAS DE ACESSO

## OBJETIVO

Desenvolver um modelo supervisionado de aprendizado de máquina capaz de classificar com acurácia medições de qualidade de Internet nas categorias de tecnologia de acesso de interesse, usando as características atreladas às medições de qualidade, à localização onde a medição é realizada e à informação do sistema de provimento de acesso no local.

## IMPLICAÇÕES SOCIOECONÔMICAS

Por meio das classificações de tecnologia de acesso preditas pelo modelo de aprendizagem de máquina, é possível realizar investigações sobre a distribuição quantitativa das tecnologias de acesso no país e sua distribuição espacial.

# PROCESSO DE SELEÇÃO DO MODELO

Os modelos de aprendizado de máquina de classificação são usados para predizer variáveis respostas categóricas e usam algoritmos para identificar regiões de não solapamento em que a mesma variável resposta é predita para todas as combinações de variáveis preditoras. Quando os modelos são "supervisionados", precisam ser treinados com bases de dados cujos elementos com atributos ou características contribuirão para definir a categoria à qual pertencem (TARCA ET AL., 2007). Em virtude de as tecnologias de acesso poderem ser caracterizadas segundo as velocidades de *download* e *upload*, o tempo de latência e a perda de pacotes, é indispensável contar com uma base de dados das métricas de qualidade da Internet no nível do setor censitário. Nesse sentido, as métricas de qualidade da Internet das medições coletadas no Brasil pelos agentes de medição do Sistema de Medição de Tráfego Internet (Medidores SIMET[1]) desenvolvidos pela equipe de Medições do CEPTRO – NIC.br[2] cumprem esses requisitos. Por exemplo, ao longo do ano de 2022, os usuários realizaram medições voluntárias de qualidade da Internet em 95% dos municípios brasileiros, distribuídas em 218 mil setores censitários[3], o que corresponde a 48% do total desses setores no Brasil. Adicionalmente, o Medidor Educação Conectada[4] monitora de forma frequente e automatizada a qualidade da Internet de aproximadamente 50.000 escolas públicas brasileiras. Essas medições são agrupadas por escola e oferecem a possibilidade de extrair estimativas estatísticas do plano contratado e da qualidade de Internet de forma mais precisa do que se fosse uma medição única por escola. Ao considerar ambos os conjuntos de dados, a informação de qualidade de Internet coletada pelo SIMET mostra-se metodologicamente viável para implementar o modelo de aprendizado de máquina e contribuir com a identificação da tecnologia de acesso, tanto nas escolas quanto nos setores censitários do Brasil.

---

[1] Mais informações disponíveis em: https://medicoes.nic.br/sobre-medicoes/

[2] Todas as figuras que ilustram este relatório foram elaboradas com base nessas medições.

[3] O setor censitário é a unidade territorial estabelecida para fins de controle cadastral, formado por área contínua, situada em um único quadro urbano ou rural, com dimensão e número de domicílios que permitam o levantamento por um recenseador. Mais informação disponíveis em IBGE (2020).

[4] Mais informações disponíveis em: https://conectividadenaeducacao.nic.br/

No processo de criação e seleção do melhor modelo, usamos a coleção de bibliotecas *TidyModels do Tidyverse* (WICKHAM ET AL. 2019) na linguagem de programação R (R Core Team 2014). Em particular, seguimos a metodologia proposta por Kuhn e Silge (2022), descrita em seguida. Para treinar e testar o modelo, estabelecemos uma base de dados em que cada observação corresponde a uma escola e cuja tecnologia de acesso à Internet é conhecida (base rotulada). Escolhemos usar as escolas para construir essa base devido à informação de qualidade de Internet existente no contexto das escolas públicas brasileiras, advindas do Medidor Educação Conectada (NIC.BR, 2023). Dois outros pontos merecem ser destacados, o primeiro diz respeito a possuir boa cobertura geográfica em extratos estaduais e municipais, e saber o tipo de localização (urbana e rural); o segundo é o fato de o medidor no contexto escolar ter uma série histórica de dados que permite extrair informação acurada sobre a estimativa do plano contratado pela escola. Para rotular a base, usamos o seguinte processo: i) obtivemos a identidade dos Sistemas Autônomos (nesse caso, delimitado aos provedores de acesso) e sua respectiva área de atuação, ao cruzar a base pública de Banda Larga Fixa da Anatel (2022), com os dados do Registro.br[5] e os dados do SIMET (NIC.BR); ii) para cada escola, identificamos o provedor de Internet e, a partir dos dados das medições SIMET, estimamos o plano de Internet contratado; iii) quando a velocidade e o provedor da escola coincidiam exclusivamente com as informações da base de provedores, atribuímos àquela escola a tecnologia declarada pelo provedor; iv) quando um mesmo provedor oferecia o mesmo plano mediante duas tecnologias diferentes, atribuímos à escola a tecnologia de maior quantidade de acessos declarados no município, segundo a base da Anatel; v) utilizamos uma terceira base de apoio para reforçar os dados rotulados, base pública do Programa Banda Larga nas Escolas[6] que possui a identificação das escolas e sua tecnologia. O uso desses dados rotulados foi utilizado somente quando, o plano de Internet, provedor e escola coincidiram com os dados do Medidor Educação Conectada, visando minimizar potenciais erros de

[5] Dados abertos disponíveis em: https://ftp.registro.br/pub/numeracao/origin/nicbr-asn-blk-latest.txt. Acesso em 15 jun. 2023

[6] Dados abertos disponíveis em: https://www.anatel.gov.br/dadosabertos/PDA/PBLE/PBLE.csv. Acesso em 15 jun. 2023.

classificação. Uma vez identificado o tipo de tecnologia usada na escola (variável resposta), adicionamos para cada escola os atributos (variáveis preditoras) que consideramos ser indicadores relevantes da presencia/ausência dos diversos tipos de tecnologia de acesso (Tabela 1).

Essa base de dados completa passou por um processo de *feature engineering*, em que eliminamos as escolas com dados faltantes para alguma das variáveis, determinamos a tipologia correta das variáveis preditoras (como categórica, numérica etc.), normalizamos as variáveis numéricas[7], verificamos a ausência de multicolinearidade entre as variáveis preditoras numéricas e mantivemos só as classes de Meio e Tecnologia de Acesso de interesse para o modelo. Como resultado de todo esse processo, obtivemos um conjunto de dados rotulado de aproximadamente 22.000 escolas. Essa base rotulada foi dividida aleatoriamente em duas partes, 70% dela usada para treinar e 30% para testar o modelo; cada uma dessa partes continha observações com proporções semelhantes de tipos de tecnologia de acesso.

O treinamento foi executado ajustando três modelos de Aprendizado de Máquina (IA): Árvores de Decisão, Floresta Aleatória e Boosted Tree (KUHN; SILGE, 2022). Com o fim de estabelecer a classificação de múltiplas classes para os modelos, indicamos o set mode de todos os modelos como "*classification*". Dessa forma, o resultado do modelo retorna a base com uma coluna indicando a categoria predita para cada escola. Para cada tipo de modelo, precisamos usar um diferente conjunto de algoritmo/biblioteca de análise: para o modelo de Árvore de decisão, usamos as ferramentas da biblioteca rpar; para o modelo de Floresta aleatória, usamos as ferramentas da biblioteca ranger; e para o modelo Boosted tree, usamos as ferramentas da biblioteca xgboost (KUHN; SILGE, 2022).

[7] Para normalizar aplicamos a seguinte fórmula aos valores das variáveis numéricas: (Valor da observação – Média desse valor) / Desvio Padrão. Dessa forma, todas as variáveis numéricas variavam de 0 a 1.

**Tabela 1.** Variáveis[8] da base de dados rotulada usada para treinar e testar os modelos

| Origem das bases de dados | Variáveis que caracterizaram as medições (por tipo) |
|---|---|
| Medições do projeto Conectividade na Educação (NIC.BR, 2023) | · Percentil 95 dos valores de velocidades de upload registrados para cada escola (Numérica; upload_percentile95)<br>· Percentil 95 dos valores de velocidades de download registrados para cada escola (Numérica; download_percentile95)<br>· Percentil 5 dos valores de latência registrados para a escola (Numérica; rtt_percentile5)<br>· Razão entre velocidade de upload e velocidade de download (Numérica; fraction_up_down_percentile95) |
| Base do Censo 2022 (INEP, 2023) | · Identidade do Setor censitário onde está localizada a escola (Numérica; "census_sector_id") |
| Dados abertos no portal da ANATEL (2022) | · Quantidade de ASN por município (Numérica; num_asn)<br>· Densidade de acessos de banda larga por cada 100 mil habitantes por município (Numérica; prop_densidade) |
| Dados sobre população e territórios de Brasil (IBGE)[9] | · Região (Categórica; gis_region)<br>· Tipo de Localização do Setor censitário (Categórica; tipo_localizacao_setor)<br>· Tipo de Localização do Município a que pertence o setor censitário (Categórica; tipo_localizacao_municip)<br>· Tamanho populacional total do município ao que pertence o setor censitário (Numérica; pop_total_unid_pop)<br>· Tamanho populacional na área densa do município ao que pertence o setor censitário (Numérica; pop_area_densa)<br>· Tamanho populacional na área densa do município ao que pertence o setor censitário (Numérica; pop_area_nao_densa)<br>· Grau de urbanização do município ao que pertence o setor censitário (Numérica; gr_urb) |

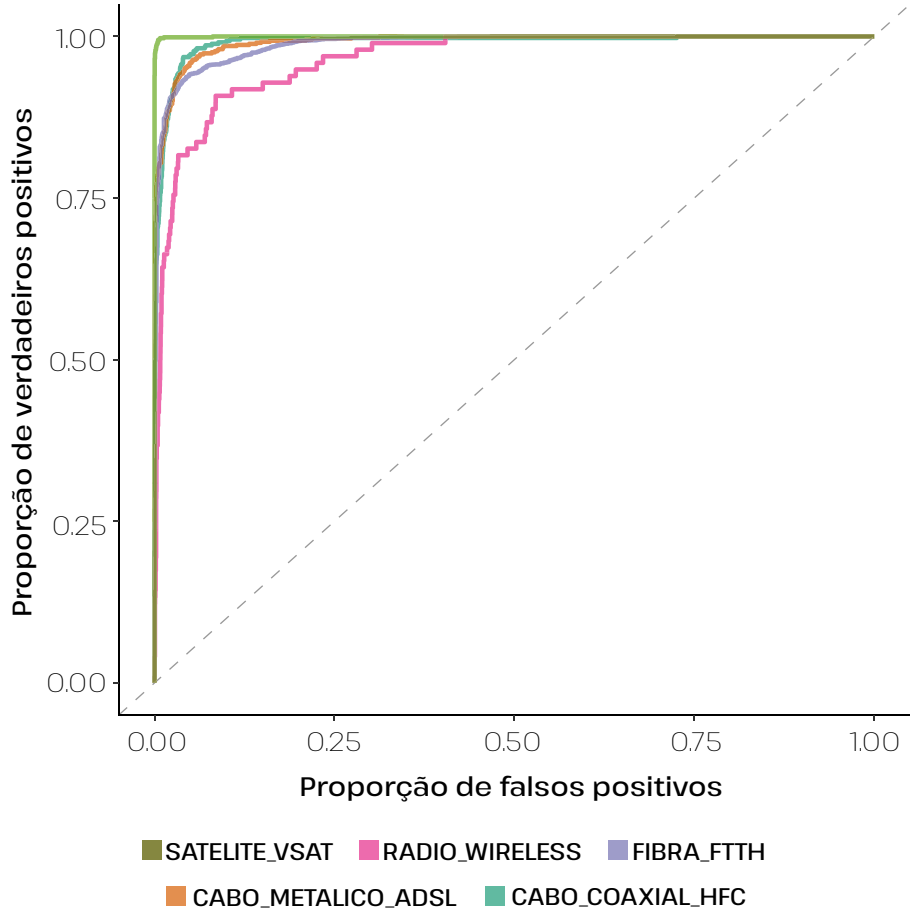
Fonte: Elaboração Própria

[8] Conforme Figura 1.

[9] Mais informações disponíveis em: https://geoftp.ibge.gov.br/organizacao_do_territorio/tipologias_do_territorio/classificacao_e_caracterizacao_dos_espacos_rurais_e_urbanos_do_brasil/. Acesso em 15 jun. 2023.

É importante explorar a performance dos modelos e de classificação ao longo de um intervalo de limiares de probabilidades para observar se o modelo consegue predizer bem de forma consistente. Para isso, testamos a performance dos três modelos de Aprendizado de Máquina usando o método de curva *Receiver Operating Characteristic* (ROC), uma análise de abordagem gráfica para analisar o desempenho de um modelo classificador (SWETS; DAWES; MONAHAN, 2000). Os valores de Area Under the ROC Curve (AUC) - que resumem a curva ROC em um único valor (entre 0 e 1) ao calcular a área sob a curva - foram: 0.88 para a Árvore de decisão, 0.99 para Floresta aleatória e 0.98 para Boosted tree. Quanto maior o AUC, maior é a acurácia do modelo em atribuir a classe tecnologia de acesso correta aos dados da base de teste; desse modo, o modelo selecionado foi o de Floresta Aleatória (Figura 1).

**Figura 1.** Resultado das curvas ROC para o modelo de Floresta Aleatória

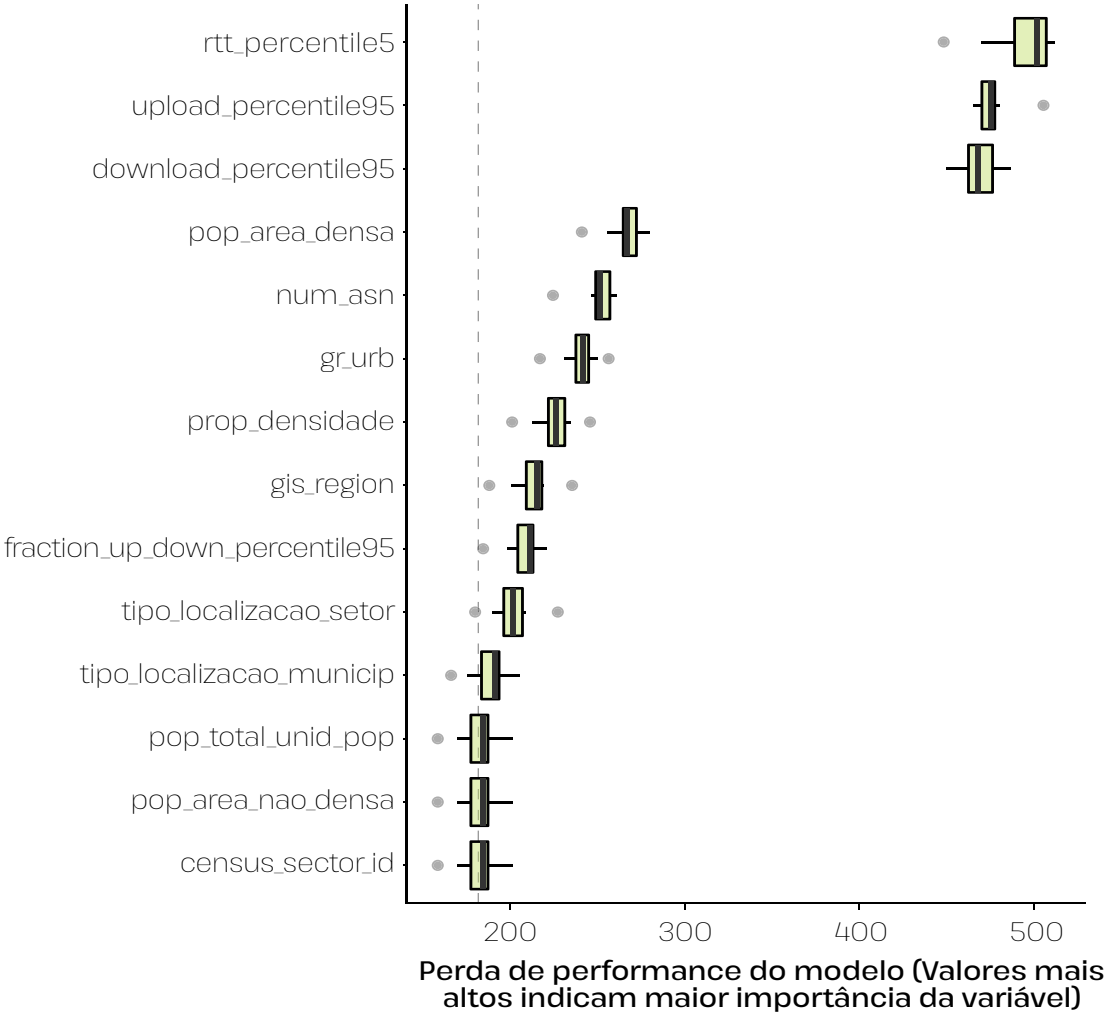



Fonte: Elaboração Própria

Cada curva ROC traça a "Proporção de verdadeiros positivos (*Sensitivity*)" *versus* a "Proporção de Falsos Positivos (1 - *Specificity*)" em diferentes limiares de probabilidade de classificação. Adicionalmente, é possível o desempenho do modelo para todas as tecnologias de acesso à Internet ao comparar as diferentes curvas. Nesse sentido, verificamos que o desempenho do modelo é melhor ao classificar Satélite do que Rádio.

Considerando o modelo de floresta aleatória, identificamos as variáveis preditoras mais importantes usando a função *explain_tidymodels* da biblioteca DALEXtra (BIECEK, 2018) (Figura 2). As três variáveis mais importantes corresponderam a dados de qualidade da conexão à Internet: velocidade de *download*, velocidade de *upload* e latência das medições.

**Figura 2.** Importância das variáveis do modelo Floresta Aleatória




Fonte: Elaboração Própria

Para identificar quais variáveis são mais importantes para influenciar a predição do modelo, permutamos os valores das variáveis (uma variável por vez): usamos o modelo para predizer e depois calculamos qual foi a perda de performance do modelo. Se a permutação de uma variável causa uma grande degradação na performance de um modelo, essa variável é importante.

# TECNOLOGIA DE ACESSO NO TERRITÓRIO

Usando o modelo selecionado, é possível realizar predições da Tecnologia de Acesso para todas as medições recebidas pelo SIMET e para as quais se sabe qual é o setor censitário desde onde a medição foi realizada. Para poder realizar essa predição, utilizamos as medições voluntárias de qualidade de Internet provenientes do medidor WEB (NIC.BR, 2023)e as combinamos com as variáveis de interesse das bases do IBGE e da ANATEL da mesma forma que foi feito para as escolas (Tabela 1). A predição resulta em estimativas da quantidade de tecnologias de acesso disponíveis (com foco nas classes: Fibra-FTTH[10], Cabo coaxial HFC[11] e outras tecnologias) por setor censitário. Embora nem todos os setores censitários tenham medições de qualidade disponíveis, os resultados modelo de aprendizado de máquina permite identificar locais (municípios ou setores censitários) com ausência ou baixa frequência de tecnologias de acesso de boa qualidade, em particular a Fibra (FTTH ou HFC), fundamental para conexões de alta velocidade e baixa latência, requisitos importantes para a utilização de uma Internet significativa em contextos domiciliares e escolares.

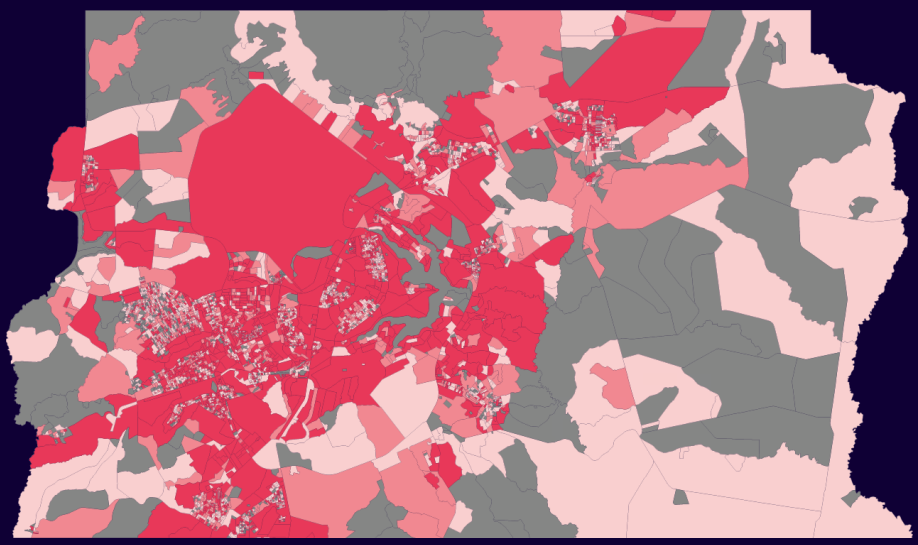
Exemplos dos resultados que podem ser obtidos por meio do modelo estão representados nas Figuras 3 e 4, que mostram o mapa da distribuição de Fibra para o Distrito Federal e a cidade de Goiânia. A variação na coloração descreve a quantidade de vezes por setor censitário em que o modelo supervisionado classificou as medições de um dado medidor como Fibra. Naqueles setores em que recebemos mais de 3 medições classificadas como Fibra a cor é mais escura, ou seja, à medida que aumenta a chance de ter Fibra no setor censitário, a cor é mais intensa. Cabe notar que, quando não há medições classifica-

[10] Fibra para a casa (*fiber to the home* – FTTH).

[11] Fibra coaxial híbrida (*hybrid fiber coax* – HFC), um tipo de conexão que utiliza cabos de fibra óptica e cabo coaxial.

das como Fibra (categoria de zero medições nos mapas), outro tipo de tecnologia foi atribuído pelo modelo, não explicitado. Já a cor cinza indica que não há informação sobre as métricas de qualidade para o setor censitário, de maneira que, segundo os dados atuais, não é possível determinar se há ou não disponibilidade de Fibra nesse território.

**Figura 3.** Resultados do modelo para os 5.150 setores censitários de Distrito Federal

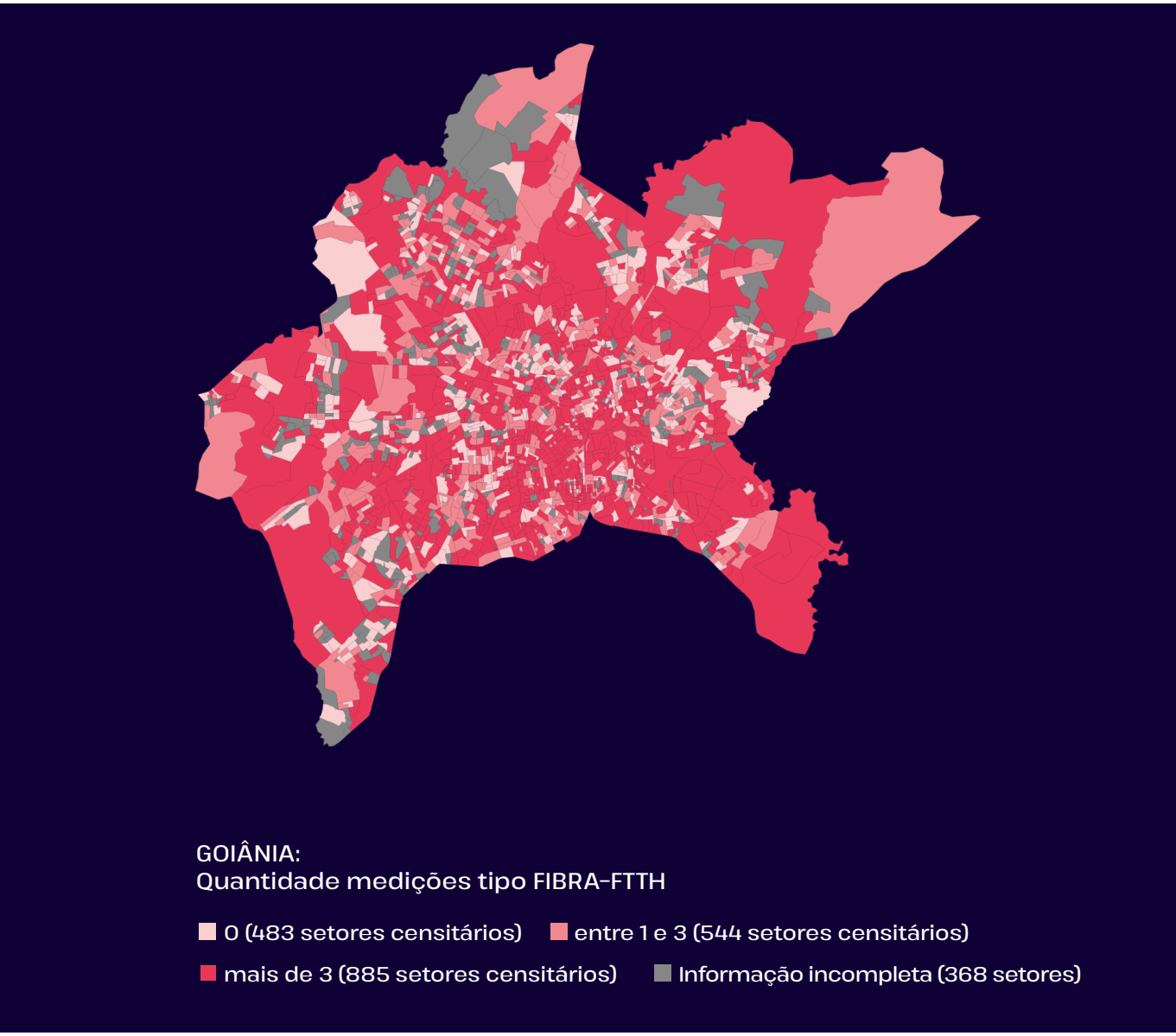



Fonte: Elaboração Própria

Classificamos o tipo de acesso à Internet para 4.222 setores censitários (82%) onde têm sido realizadas medições de qualidade nos últimos 2 anos. Em 2.748 setores censitários, há evidência de disponibilidade de acesso à Internet via Fibra, ou seja, o modelo classificou pelo menos uma medição como Fibra.

Nas áreas com cor vermelha mais clara - que correspondem a 35% dos setores censitários estudados -, não há registro de acesso à Internet via Fibra, portanto são localidades com baixa chance de haver Fibra cuja disponibilidade de acesso a uma Internet de qualidade significativa é reduzida.

**Figura 4.** Resultados do modelo para os 2.280 setores censitários de Goiânia



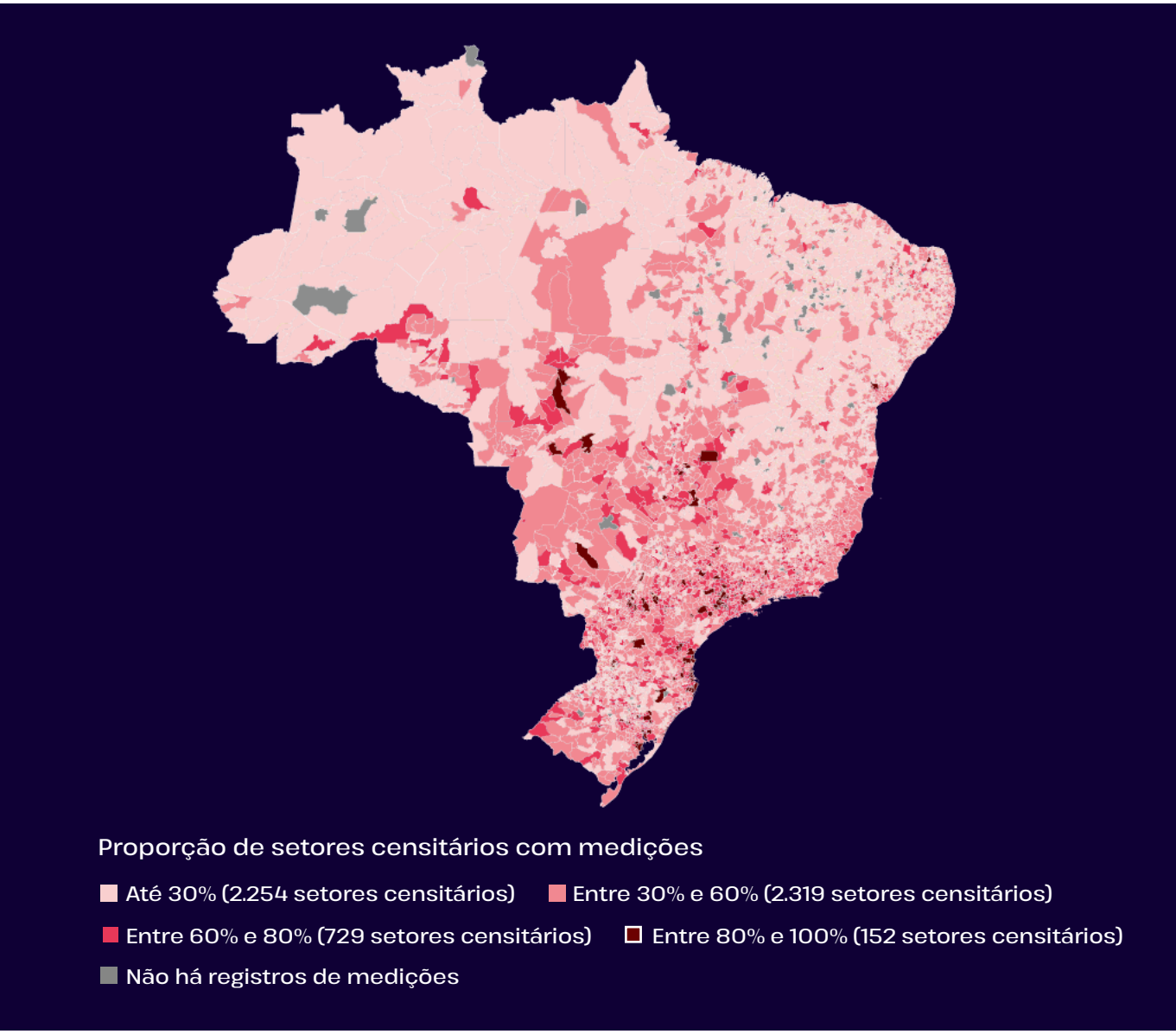

Fonte: Elaboração Própria

Classificamos o tipo de acesso à Internet para 1.912 setores censitários (84%) onde têm sido realizadas medições de qualidade nos últimos 2 anos. Em 1.429 setores censitários, há evidência de disponibilidade de acesso à Internet via Fibra, ou seja, o modelo classificou pelo menos uma medição como Fibra. Nas áreas com cor vermelha mais clara - que correspondem a 25% dos setores censitários estudados -, não há registro de acesso à Internet via Fibra, portanto são localidades com baixa chance de ter Fibra cuja disponibilidade de acesso a uma Internet de qualidade significativa é reduzida.

Nos territórios, em que uma qualidade de Internet significativa está disponível, é importante melhorar as estratégias que permitam aos usuários reconhecer as vantagens dessa tecnologia e manter o custo acessível. Já nos territórios onde a probabilidade de ter acesso a uma Internet de qualidade significativa é reduzida, deve haver políticas públicas que incentivem o investimento massivo para expansão de fibra nessas regiões.

É importante ressaltar que outras tecnologias e outros níveis territoriais podem ser mapeados usando a metodologia descrita. Por fim, assim como todos os modelos de aprendizagem de máquina supervisionado, a performance do modelo depende dos dados usados para realizar a previsão. Nesse caso, como as métricas de qualidade são essenciais para fazer a inferência, o modelo é particularmente útil nas regiões onde há essa informação.

**Figura 5.** Mapa da proporção atual de setores censitários com registo de medições de qualidade para cada um dos municípios do Brasil



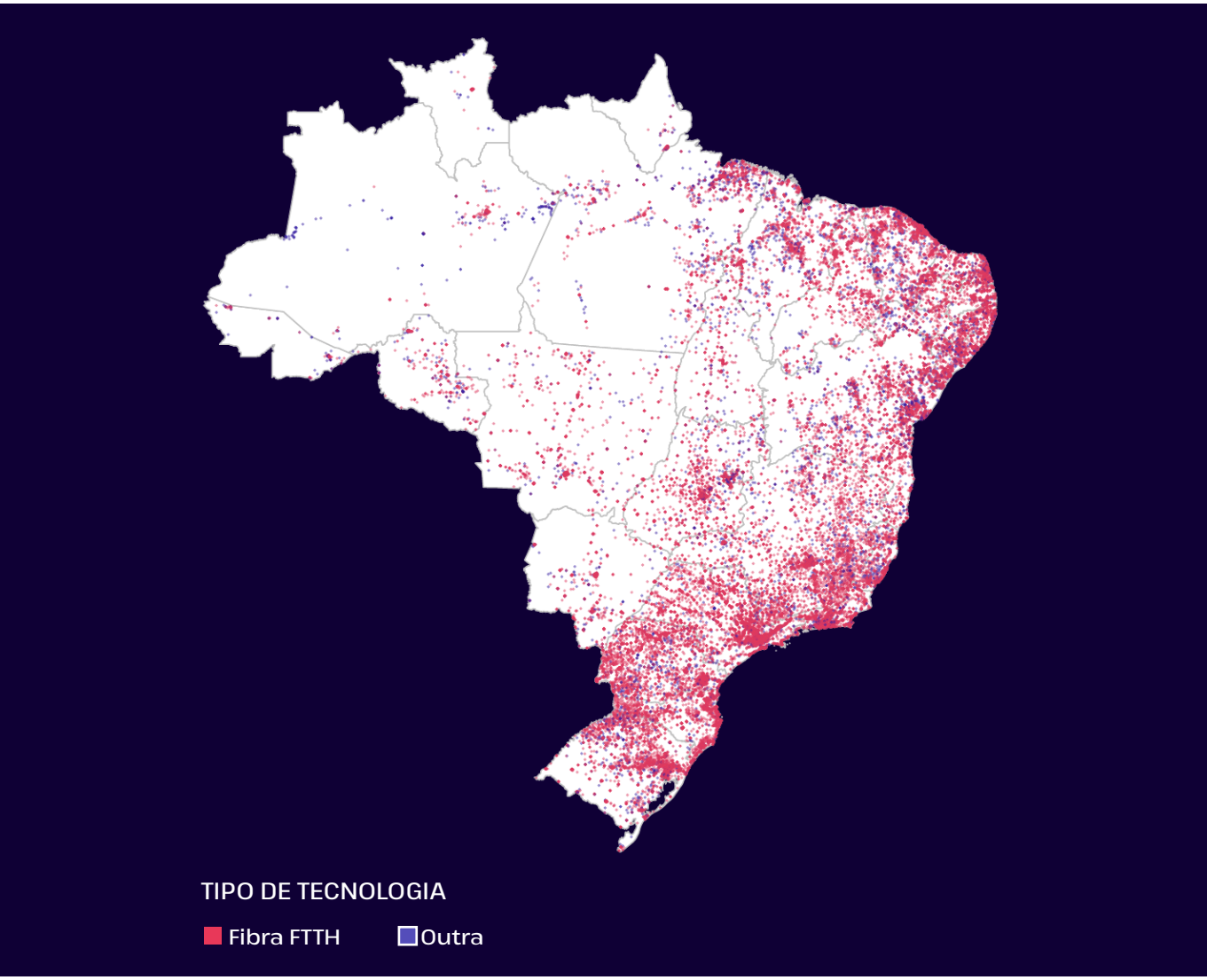

Fonte: Elaboração Própria

À medida que a cor é mais escura, aumenta a probabilidade de o modelo conseguir predizer a tecnologia de acesso em escala inframunicipal. Nesse caso, foram usados os dados das medições registradas entre janeiro de 2021 e março de 2023.

# TECNOLOGIA DE ACESSO DAS ESCOLAS

O modelo inferencial treinado também pode ser usado em projetos especiais, como é o monitoramento da conectividade na educação. Nesse contexto, podemos usá-lo para identificar a tecnologia de acesso para todas as escolas públicas em atividade do Brasil, das quais o Medidor Educação Conectada recebeu dados de qualidade da conexão à Internet no último semestre (Figura 6).

**Figura 6.** Localização das 56.394 escolas com medidor Educação Conectada Instalado e informação sobre a tecnologia de acesso




Fonte: Elaboração Própria

Para facilitar a visualização criamos unicamente duas categorias: a escola tem Fibra ou a escola tem um tipo de tecnologia diferente de Fibra.

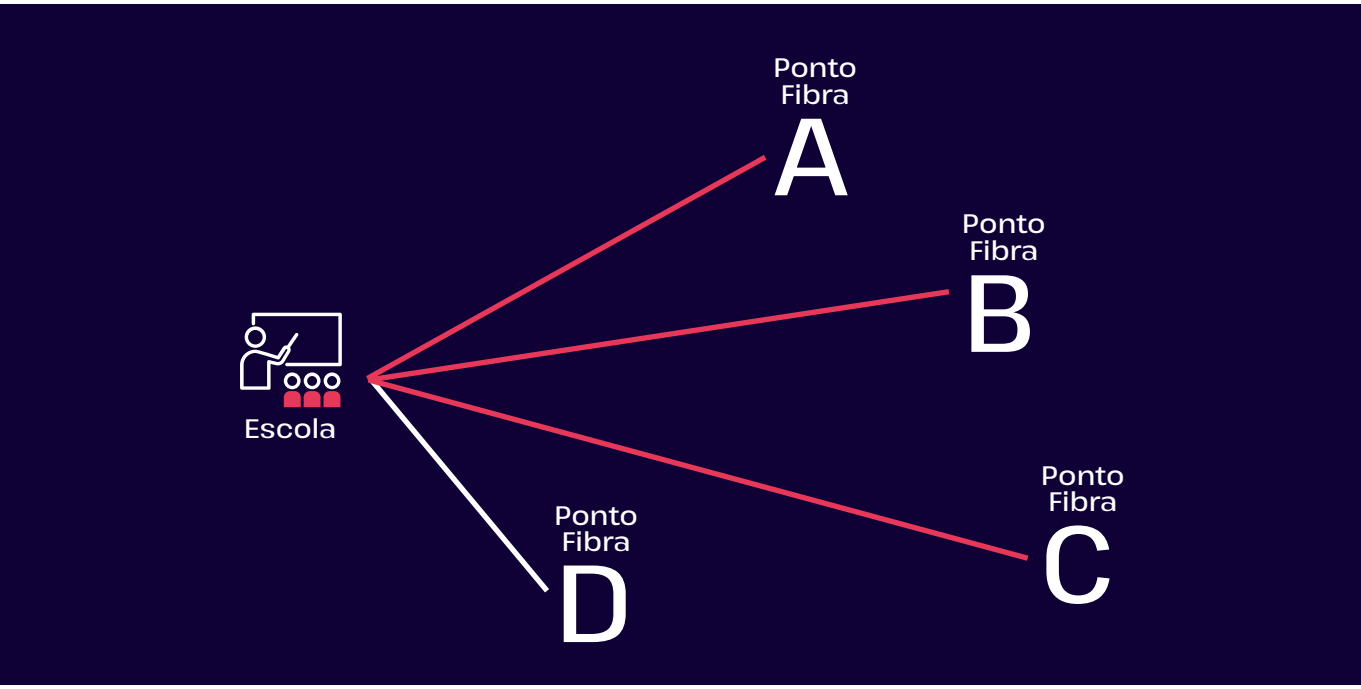
Adicionalmente, a fim de considerar também as escolas que não têm se incorporado ainda ao projeto e para as quais não temos dados da qualidade da Internet, foi necessário incluir as informações sobre tecnologia de acesso advindas dos territórios adjacentes às escolas. Nesse contexto, seguimos o seguinte protocolo: i) usamos as predições da Tecnologia de Acesso para todas as medições georreferenciadas recebidas pelos medidores SIMET ao longo do último ano; ii) mantivemos só as medições de Internet identificadas como Fibra-FTTH ou Cabo HFC; iii) excluímos as escolas sem Fibra (segundo o modelo); iv) usando a função *distm* biblioteca *geosphere* (HIJMANS, 2022), criamos uma matriz de distância (em km) entre o conjunto de localizações das medições de Fibra e o conjunto de localizações das escolas selecionadas; v) para cada escola, identificamos a medição de Fibra mais próxima e explicitamos a distância que separa as localizações (Figura 7); iv) usamos a informação sobre a presença ou ausência de Fibra e, em conjunto com as informações da distância mínima ao ponto de Fibra mais próximo, incluímos as escolas em alguma das seguintes categorias:

**1.** Acesso adequado: Escolas com Fibra ou até 1 km de um lugar em que foi identificada uma medição com Fibra.

**2.** Acesso inadequado - Fibra próxima: Escolas sem Fibra ou até 20 km de um lugar em que foi identificada uma medição com Fibra.

**3.** Acesso inadequado - Fibra distante: Escolas sem Fibra com mais de 20 km de um lugar em que foi identificada uma medição com Fibra ou de localização imprecisa (foi assumido que escolas sem localização precisa estão em territórios bastante distantes de centros urbanos).

**Figura 7.** Visualização do processo de cálculo de distância entre a escola e o ponto de Fibra




Fonte: Elaboração Própria

Calculamos a distância entre a localização da escola e a localização das medições que foram classificadas como Fibra pelo modelo. O ponto de fibra mais próximo (Ponto de Fibra D) foi selecionado: a distância mínima até o ponto de Fibra mais próximo foi usada para atribuir uma categoria de conectividade às escolas, as quais servirão para orientar a estratégia de conectividade mais apropriada para cada escola, a fim de garantir o custeio para a contratação de planos de Fibra. Desse modo, investe-se na expansão de infraestrutura desse modelo ou, em caso de escolas bem distantes de um ponto de Fibra, garantindo o custeio de planos de tecnologias de acesso à Internet alternativas à Fibra ótica que sejam apropriadas para levar uma conectividade significativa, a fim de permitir o desenvolvimento das competências digitais nas escolas públicas brasileiras (BETTEGA; MARIN; KUESTER, 2020).

Adicionalmente, cabe ressaltar que, para poder contribuir na conectividade das escolas, é importante que as escolas façam uso do Medidor Educação Conectada e acessem a

plataforma Conectividade na Educação para acompanhar as medições de qualidade de Internet. Por fim, é essencial ter conhecimento da localização precisa das escolas: no momento, são 1.430 escolas públicas em atividade para as quais a localização é imprecisa.

# CONCLUSÃO

Ambas as abordagens presentadas podem apoiar os formuladores de políticas públicas e a sociedade em geral na elaboração de um diagnóstico amplo e preciso da distribuição dos diferentes tipos de tecnologias de acesso à Internet em todo o país e, assim, ajudar a reduzir a desigualdade de acesso à Internet mediante o apoio ao uso de tecnologias de acesso modernas que proporcionem conectividade significativa.

# REFERÊNCIAS


AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES (ANATEL). *Banda Larga Fixa*. Brasília: Ministério das Comunicações, 2022. Disponível em: https://informacoes.anatel.gov.br/paineis/acessos/banda-larga-fixa. Acesso em 25 maio 2023.

BETTEGA, E., MARIN, G., KUESTER, P. Os limites da banda larga: o papel da conectividade nos usos das TIC para o desenvolvimento das competências digitais nas escolas públicas brasileiras. *In*: NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR (NIC.BR). *Pesquisa sobre o uso das tecnologias de informação e comunicação nas escolas brasileiras*: TIC Educação 2019. São Paulo: CGI.br, 2020. p. 147-162. 2020. Disponível em: https://cetic.br/media/docs/publicacoes/2/20201123090444/tic_edu_2019_livro_eletronico.pdf. Acesso em 23 maio 2023.

BIECEK, P. DALEX: Explainers fo r complex predictive models in R. *Journal of Machine Learning Research*, v. 19, n. 84, p. 1-5, 2018. Disponível em: https://jmlr.org/papers/volume19/18-416/18-416.pdf. Acesso em 23 maio 2023.

HIJMANS, R. Geosphere: Spherical Trigonometry. R package version 1.5-18. *The Comprehensive R Archive Network (CRAN)*, 15 nov. 2022. Disponível em: https://CRAN.R-project.org/package=geosphere. Acesso em 23 maio 2023.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). *Malha de Setores Censitários*. IBGE: 2020. Disponível em: https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/26565-malhas-de-setores-censitarios-divisoes-intramunicipais.html?edicao=30113&t=saiba-mais-edicao. Acesso em 25 maio 2023.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). *Censo Escolar 2022*. Brasília: INEP, 2023. Disponível em: https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar/resultados/2022. Acesso em 25 maio 2023.

KUHN, M.; SILGE. *Tidy Modeling With R*: A Framework for Modeling in the Tidyverse. Sebastopol: O´Reilly Media, 2022.

NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR (NIC.br). *Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros*: TIC Domicílios 2022 (Microdados). São Paulo: CGI.BR, 2023. Disponível em: https://cetic.br/pt/arquivos/domicilios/2022/individuos/. Acesso em 23 maio 2023.

SWETS, J.; DAWES, R.; MONAHAN, J. Better decisions through science. *Scientific American*, v. 283, n. 4, p. 82-87, out. 2000. Disponível em: https://pubmed.ncbi.nlm.nih.gov/11011389/. Acesso em 23 maio 2023.

TARCA, A. L. ET AL. Machine learning and its applications to biology. *PLoS computational biology*, v. 3, n. 6, e116, 29 jun. 2007. Disponível em: https://doi.org/10.1371/journal.pcbi.0030116. Acesso em 23 maio 2023.

WICKHAM, H. ET AL. Welcome to the Tidyverse. *The Journal of Open Source Software*, v. 4, n. 43, art. 1686, 21 nov. 2019. Disponível em: https://doi.org/10.21105/joss.01686 . Acesso em 15 jun. 2023.

ceptro.br nic.br cgi.br