



## RELAÇÕES ENTRE ESTADO NUTRICIONAL DE PRÉ-ESCOLARES E INDICADORES SOCIOECONÔMICOS E DE INFRAESTRUTURA URBANA: UMA ABORDAGEM BASEADA EM ALGORITMO CART

**T. F. Olivatto, C. S. Pitombo, A. L. Cunha e E. A. Melanda**

### RESUMO

O estado nutricional frequentemente subsidia a tomada de decisão do poder público, sendo, tradicionalmente, classificado a partir do Índice de Massa Corporal. Contudo, estudos envolvendo indicadores socioeconômicos e de infraestrutura urbana vêm se mostrando promissores. Objetivando compreender as relações destes indicadores com estado nutricional, foi conduzido um estudo com base no algoritmo *Classification And Regression Tree* (CART) referente a dados nutricionais de pré-escolares de Araraquara-SP. As variáveis socioeconômicas que melhor explicam a variabilidade dos dados são “número de pessoas” e “escolaridade”, sendo que nas famílias mais numerosas e com menor nível de escolaridade houve maior ocorrência de desnutrição. Dentre as variáveis de infraestrutura se destacam “calçada” e “padrão da casa”, verificando-se que a presença de calçada remete a maior ocorrência de eutrofia, e melhores residências, a maior ocorrência de obesidade. O uso do algoritmo CART auxiliou a investigação das relações entre os indicadores estudados e a ocorrência dos estados nutricionais.

### 1 INTRODUÇÃO

De acordo com o relatório produzido pela Organização Mundial da Saúde (OMS), uma a cada três pessoas em todo o mundo sofre com desnutrição, deficiência de vitaminas e minerais, excesso de peso, obesidade ou doenças não transmissíveis relacionadas à dieta. Este mesmo relatório estimou que, em 2019, 38 milhões de crianças menores de 5 anos apresentavam sobre peso ou obesidade enquanto 61,3 milhões estavam abaixo do peso e 144 milhões gravemente desnutridas (WHO, 2020).

Neste contexto, a governança nutricional se revela um importante tópico no ambiente político, com foco principal em planos e medidas para promover uma nutrição mais saudável. No que diz respeito a iniciativas públicas e medidas regulatórias, destacam-se a rotulagem nutricional, políticas de incentivo à atividades fiscais, implementação de espaços públicos para desenvolvimento destas atividades, controle de gorduras trans e reformulação de produtos alimentícios, principalmente os voltados para crianças (WHO, 2020).

Segundo Fajersztajn *et al.* (2016), doenças associadas à comportamentos nutricionais, como diabetes tipo 2, hipertensão arterial e obesidade estão diretamente relacionadas ao estilo de vida urbano. Isso se deve, principalmente, à forma como o ambiente urbano se apresenta, podendo contribuir ou inviabilizar hábitos individuais saudáveis, como por exemplo, a prática de atividade física em espaços públicos adequados e bem distribuídos ao longo do território. Sendo assim, indicadores nutricionais podem ser utilizados como ferramentas adicionais na gestão integrada de políticas da saúde coletiva e de urbanismo.

Neste sentido, visando nortear a tomada de decisão, destacam-se os planos de desenvolvimento relacionados à nutrição. O principal deles é o Plano sobre Nutrição Materna, Infantil e Juvenil (MIYCN, em inglês *Maternal, Infant and Young Child Nutrition*) resultante da 65<sup>a</sup> Assembleia Mundial da Saúde, ocorrida em maio de 2012. Das seis metas globais estabelecidas neste plano, três delas se referem à redução de ocorrência de atrofia em crianças menores de 5 anos, bem como redução de casos de crianças obesas e desnutridas (WHO, 2014).

O desenvolvimento de planos e medidas é baseado em diversos indicadores e estudos. Mundialmente, os indicadores mais aceitos e utilizados são os estabelecidos pela OMS. Estes indicadores focam principalmente em ocorrência de atrofia, anemia, nascimentos abaixo do peso, obesidade, amamentação e desnutrição (WHO, 2014). Tradicionalmente, a partir da relação de peso e altura de um indivíduo, calcula-se o seu IMC (Índice de Massa Corporal), que então é utilizado para estabelecer o estado nutricional de um indivíduo que classifica-se em: desnutrição, eutrofia (adequado), sobrepeso e obesidade (Barbosa Filho *et al.*, 2010).

Além desta classificação, alguns estudos incluem indicadores populacionais e socioeconômicos para a determinação do estado nutricional infantil (Foschini e Campos, 2010; WHO, 2014). O estudo desenvolvido por Foschini e Campos (2010), em Araraquara-SP (Brasil), relaciona o estado nutricional infantil a aspectos como nível de escolaridade e nível econômico da família, sexo e número de pessoas que residem com a criança.

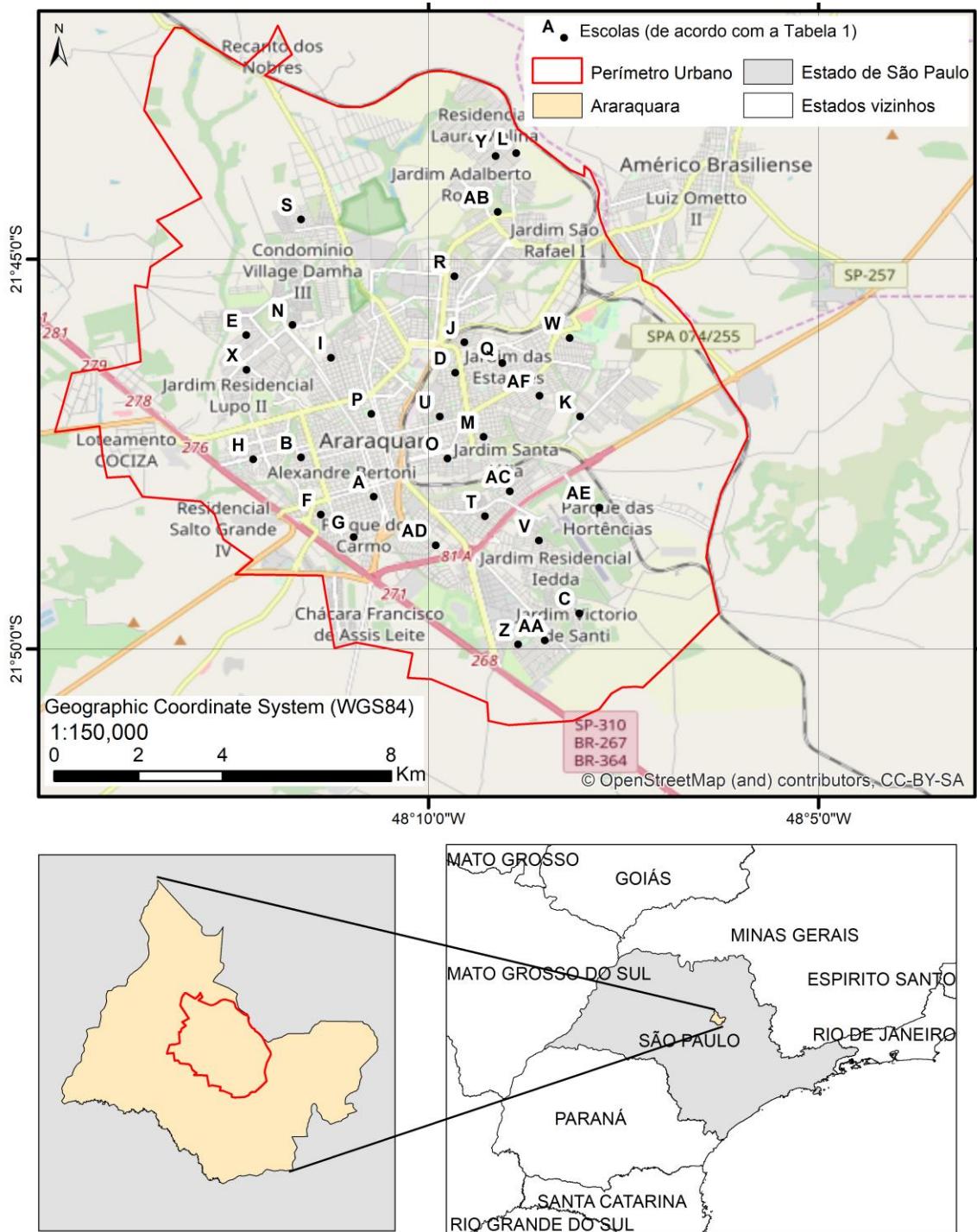
Apesar de considerar aspectos como renda e número de pessoas na residência, estudos que incluam variáveis de infraestrutura urbana às quais as crianças estão sujeitas são escassos. Barcellos e Ramalho (2002) afirmam que, para compreender questões de saúde e seus determinantes, deve-se considerar além de indicadores tradicionais de saúde, indicadores referentes à população, incluindo dados que considerem o espaço em que esta população está inserida. As características de ambiente e estruturas podem ser vistas como convergência dos processos envolvidos na determinação da qualidade de vida das populações. Um exemplo da importância da inclusão de indicadores ambientais, são os aspectos de infraestrutura urbana de acesso a serviços de saneamento considerados pela OMS para estudos nutricionais (WHO, 2014).

Deste modo, reconhecendo a escassez de estudos que incluam indicadores alternativos, o objetivo deste trabalho é realizar uma análise exploratória de relações entre o estado nutricional de pré-escolares e indicadores socioeconômicos e de infraestrutura urbana. Para tal, utilizou-se como ferramenta principal a classificação por árvore de decisão, com base no algoritmo CART (*Classification and Regression Tree*).

## 2 MATERIAIS E MÉTODOS

### 2.1 Materiais

O estudo foi desenvolvido no perímetro urbano do município de Araraquara (área ilustrada pela Figura 1). Este município está localizado na região central do estado de São Paulo, no Brasil, estando à aproximadamente 273 km da capital São Paulo. A população estimada é de 238.339 habitantes, com uma densidade demográfica de 237,62 hab./km<sup>2</sup> (IBGE, 2020).



**Figura 1. Localização da área de estudo, incluindo as escolas envolvidas neste estudo**

O Índice de Desenvolvimento Humano (IDH) do município é de 0,815 e se enquadra na faixa de IDH muito alto (entre 0,800 e 1). Deste índice, a dimensão que mais contribui é Longevidade, com índice de 0,877, seguida de Renda, com índice de 0,788, e de Educação, com índice de 0,782. Em 2010, a mortalidade de crianças com menos de 5 anos de idade no município foi de 11,9 óbitos por mil nascidos vivos. Em 1991, era de 25,7 óbitos por mil nascidos vivos (PNUD *et al.*, 2010). A dimensão de Longevidade do IDH e a mortalidade tem relação direta com saúde e qualidade de vida da população e, indireta, com estado nutricional.

O município conta com 43 Centros de Educação e Recreação (CER) (Prefeitura Municipal de Araraquara, 2020), contudo o banco de dados obtido representa apenas 32 deles. Para o desenvolvimento do método proposto, os nomes dos CERs foram associados a letras, objetivando viabilizar a visualização dos mesmos na árvore de decisão (AD) e no mapa resultante. A Tabela 1 associa o nome do CER à simbologia (letra) utilizada e a Figura 1 apresenta as suas localizações.

**Tabela 1. Relação de CERs e simbologia associada**

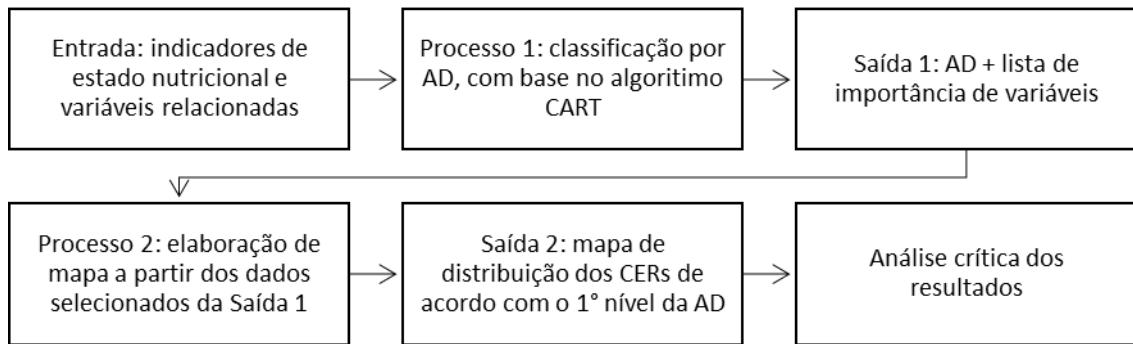
Nome do CER	Simbologia associada	Nome do CER	Simbologia associada
Carmelita Garcez	A	Álvaro W. Colino	Q
Eduardo B. Coelho	B	Honorina C. Lia	R
Antonio C. de Lima	C	Maria da Gloria F. Simões	S
Jose do Amaral Velosa	D	Amelia F. Manini	T
Maria Enaura M. Magalhaes	E	Eloa do V. Quadros	U
Concheta S. Mendonca	F	Marialice L. Tedde	V
Pe. Bernardo Plate	G	Antonio T. P. Lima	W
Jacomina F. Sambiase	H	Ricardo C. Monteiro	X
Cyro G. Ramos	I	CAIC Prefeito Rubens Cruz	Y
Eunice B. T. Piza	J	Maria Renata Lupo	Z
Eudoxia P. Ferraz	K	Maria José Porciuncula	AA
Judith de B. Batelli	L	Maria Pradelli Malara	AB
Cotinha de Barros	M	José Pizani	AC
Zida M. Pierri	N	Maria Barcarolla Filie	AD
Adelina L. Amaral	O	Anunciata Lia David	AE
Leonor M. de Barros	P	Rosa R. Stringuetti	AF

Ao todo foram considerados os dados de 491 pré-escolares (crianças de zero a cinco anos, 11 meses e 29 dias de idade) de 32 CERs. Para preservar a característica de anonimato dos dados, cabe ressaltar que não se teve acesso aos nomes dos pré-escolares e nem aos seus respectivos endereços (os dados de cada criança foram associados à escola que o mesmo frequenta). Os dados foram obtidos durante pesquisa realizada por Foschini e Campos (2010), na qual o levantamento de dados foi aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Ciências Farmacêuticas da UNESP (protocolo: 15/2006).

Para construção da AD, os dados foram importados para o software *IBM SPSS Statistics 20* e foi utilizado o algoritmo CART (*Classification and Regression Tree*). Este software permite, além da visualização da AD, a observação de outros parâmetros relevantes, como por exemplo, a relação das variáveis independentes que mais influenciam a classificação. Para a elaboração dos mapas foi utilizado o software livre QGis.

## 2.2 Método

O Fluxograma da Figura 2 apresenta as principais etapas envolvidas no método desenvolvido.



**Figura 2. Fluxograma das etapas do método**

As variáveis utilizadas nesta investigação, bem como seus atributos e descrição, estão delineadas na Tabela 2. Foi adotada como variável dependente o Estado Nutricional dos pré-escolares dos CERs, e as outras variáveis foram as variáveis independentes.

Neste trabalho, optou-se pela ferramenta de Árvore de Decisão (AD), por ser considerada uma técnica simplificada de mineração de dados capaz de representar as relações de determinado conjunto de dados (Medeiros *et al.*, 2014; Quilan, 1983). Nesta técnica, a classificação dos dados se dá por meio de algoritmos, dentre os quais os mais utilizados são o C4.5 (Quilan, 1983), CHAID (Kass, 1980) e CART (Breiman *et al.*, 1984).

A repetição do procedimento de decisão ocorre em vários níveis, optando-se para o desenvolvimento deste trabalho o máximo de 5 níveis, conforme limitação do *software* utilizado. Os algoritmos também atuam com regras de parada (*stopping rules*), que no caso do CART corresponde ao número de mínimo de observações nos nós terminais (Breiman *et al.*, 1984), as quais foram de 5 observações nas análises realizadas.

O algoritmo CART utiliza o *índice Gini* para a partição binária dos dados, sendo que este pode variar de 0 a 1. Quanto mais próximo de 1, mais impuro é um nó. O objetivo do algoritmo é tornar os nós cada vez mais homogêneos (*Índice Gini* mais próximo ao valor zero). Consequentemente, o critério *Gini* ramifica as observações a partir dos atributos das variáveis independentes, considerando a maior ocorrência de uma única categoria da variável dependente (Breiman *et al.*, 1984; Pitombo, *et al.*, 2017; Souza *et al.*, 2017). Neste caso, ponderando que o foco foi verificar as ocorrências de desnutrição, eutrofia, sobre peso e obesidade, cada bipartição ocorrerá a partir do valor, ou intervalo de valor, de determinada variável independente que resultar nos nós mais puros segundo categoria de estado nutricional.

Além de observar padrões estatísticos, compreender a distribuição espacial dos fenômenos constitui uma importante ferramenta para subsidiar conhecimentos nas mais diversas áreas, incluindo a área da saúde (Druck *et al.*, 2004). Desta forma, objetivando também compreender melhor a distribuição espacial dos CERs segundo características nutricionais, foi elaborado um mapa com a discriminação dos CER relacionados nos dois nós filhos resultantes do 1º nível da AD.

**Tabela 2. Variáveis, atributos e descrição**

Variável	Atributos	Descrição
Estado Nutricional	Desnutrição (estado nutricional inadequado) Eutrofia (estado nutricional adequado) Sobrepeso (estado nutricional inadequado) Obesidade (estado nutricional inadequado)	Estado Nutricional do pré-escolar de acordo com as faixas de IMC por faixa etária de Sellen (1998).
Escola	de A a AF	Nomenclatura dos CERs de acordo com a Tabela 1.
Número de pessoas	de 2 a 10	Número de pessoas residentes com o pré-escolar.
Sexo	Feminino Masculino	Declarado.
Nível Econômico	1 - A1 2 - A2 3 - B1 4 - B2 5 - C 6 - D 7 - E	Nível Econômico de acordo com os parâmetros estabelecidos pelo Critério de classificação econômica Brasil (ABEP, 2008). Esta classificação varia de A1, classe de maior renda, à E, classe de menor renda.
Escolaridade	1 - Analfabeto ou primário incompleto 2 - Primário completo ou ginásial incompleto 3 - Ginásial completo ou colegial incompleto 4 - Colegial completo ou superior incompleto 5 - Universitário completo	Escolaridade do membro da família com maior grau de instrução.
Asfalto	0 - Sem 1 - Ruim 2 - Bom	Percepção visual.
Guia/Sarjeta	Sem Com	Percepção visual.
Muro/Grade	Sem Com	Referente à residência do pré-escolar. Percepção visual.
Calçada	0 - Sem 1 - Ruim 2 - Boa 3 - Bonita	Percepção visual.
Recuo	Sem Com	Referente à residência do pré-escolar. Percepção visual.
Padrão da Casa	0 - Péssimo 1 - Ruim 2 - Bom 3 - Ótimo	Percepção visual.
Rio (confrontação)	Sim Não	Percepção visual. Confrontação da residência do pré-escolar com córrego.
Coleta de Esgoto	Sim Não	Percepção visual.

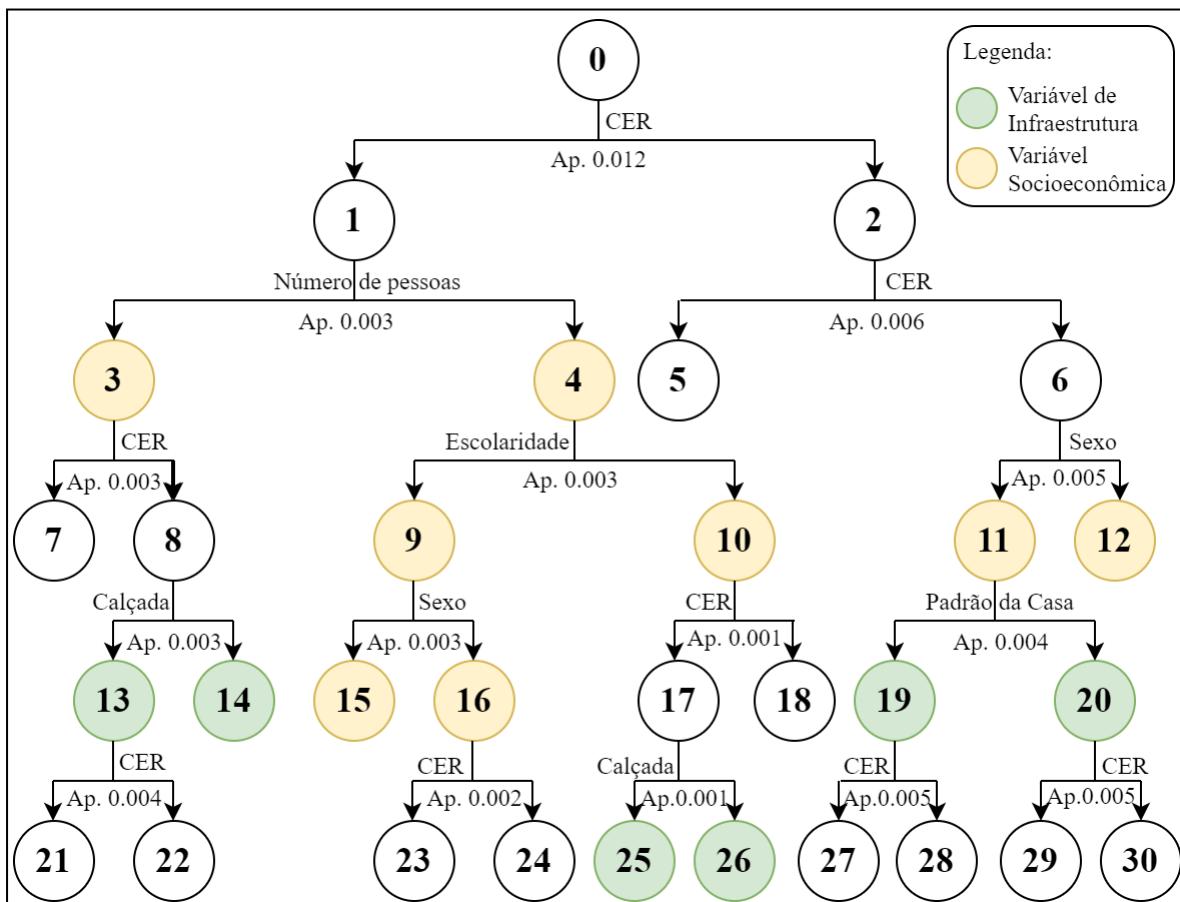
Fonte: Adaptado de Foschini e Campos (2010).

### 3 RESULTADOS E DISCUSSÕES

A visualização da distribuição da ocorrência dos estados nutricionais no município, de acordo com a instituição educacional a qual o pré-escolar frequenta, pode ser uma importante ferramenta de gestão pública, auxiliando a tomada de decisão relacionada a questões de saúde e de infraestrutura do município. Portanto, as análises realizadas podem auxiliar na compreensão do relacionamento de variáveis socioeconômicas e de infraestrutura urbana com a ocorrência de estados nutricionais. Neste aspecto, a análise de dados multivariados, a partir da técnica mineração de dados, foi essencial na extração dos

possíveis padrões desconhecidos previamente ou relações entre as diversas variáveis envolvidas.

A Figura 3 é a representação do gráfico acíclico do algoritmo CART resultante. De forma geral, a estrutura gráfica resultante representa as hierarquias de escolhas das variáveis explicativas, sendo que em cada nó foi obtida uma classe relativamente homogênea - de acordo com o *índice Gini* e *aprimoramento* identificado (vide “Ap.” na Figura 3). Nesta figura, os nós resultantes de variáveis explicativas socioeconômicas e de infraestrutura estão identificados em cores diferentes.



**Figura 3. Esquema da Árvore de Decisão (algoritmo CART)**

A Tabela 3 detalha os nós representados na Figura 3, permitindo a identificação da variável explicativa responsável pela divisão, os atributos contidos nos agrupamentos resultantes destas divisões e as porcentagens de ocorrência dos estados nutricionais em cada nó. As linhas referentes aos nós terminais estão destacadas em cinza na Tabela 3.

Os nós resultantes nos diferentes níveis são caracterizados por valores das variáveis explicativas (independentes), as quais apresentem maior percentual de frequência dos estados nutricionais (eutrofia, desnutrição, sobre peso e obesidade). Consequentemente, pode-se identificar quais fatores influenciam a predominância de ocorrência de determinado estado nutricional em relação aos outros.

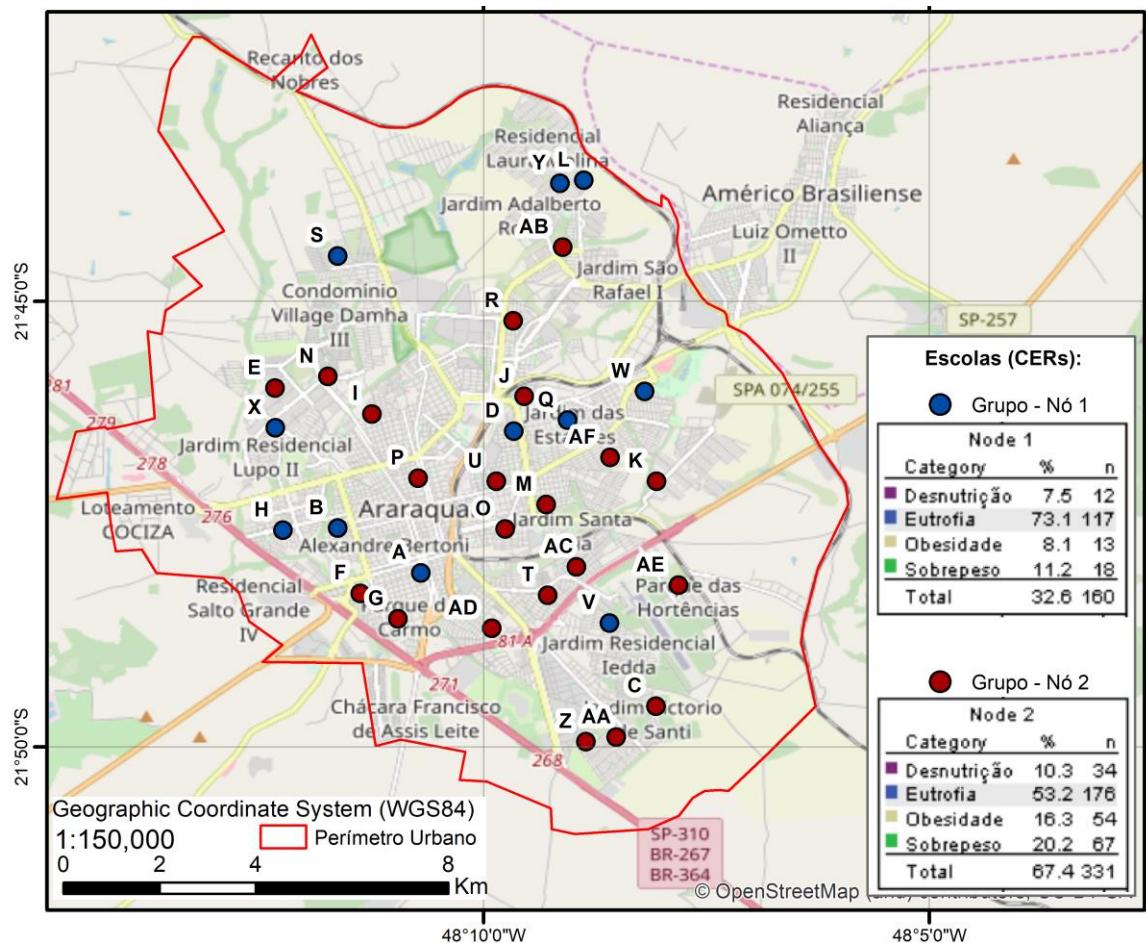
**Tabela 3. Relação dos nós e divisões da Árvore de Decisão**

Nó	Variável explicativa	Agrupamento Resultante	D(%)	E(%)	O(%)	S(%)
0	Estado Nutricional	(nó raiz)	9,4	59,7	13,6	17,3
1	CER	A, B, D, H, L, Q, S, V, W, X, Y	7,5	73,1	8,1	11,2
2	CER	C, E, F, G, I, J, K, M, N, O, P, R, T, U, Z, AA, AB, AC, AE, AF	10,3	53,2	16,3	20,2
3	Número de pessoas	<=4,5	3,7	70,1	12,1	14,0
4	Número de pessoas	>4,5	15,1	79,2	0,0	5,7
5	CER	C, E, F, G, J, U, Z, AB, AC, AD	15,4	51,7	8,4	24,5
6	CER	I, K, M, N, O, P, R, T, AA, AE, AF	6,4	54,3	22,3	17,0
7	CER	D, L, W	3,1	65,6	3,1	28,1
8	CER	A, H, Q, S, V, X, Y	4,0	72,0	16,0	8,0
9	Escolaridade	<=2	29,2	66,7	0,0	4,2
10	Escolaridade	>2	3,4	89,7	0,0	6,9
11	Sexo	Feminino	8,3	48,1	29,6	13,9
12	Sexo	Masculino	3,8	62,5	12,5	21,2
13	Calçada	<=0 (sem calçada)	6,2	50,0	25,0	18,8
14	Calçada	>0 (com calçada)	3,4	78,0	13,6	5,1
15	Sexo	Feminino	50,0	37,5	0,0	12,5
16	Sexo	Masculino	18,8	81,2	0,0	0,0
17	CER	B, L, S	7,7	76,9	0,0	15,4
18	CER	A, D, H, Q, V, W, X, Y	0,0	100	0,0	0,0
19	Padrão da Casa	<=1 (ruim e péssimo)	7,6	58,1	22,7	13,6
20	Padrão da Casa	>1 (bom e ótimo)	9,5	35,7	40,5	14,3
21	CER	S, V	14,3	14,3	42,9	28,6
22	CER	Q, X, Y	0,0	77,8	11,1	11,1
23	CER	L, S, V	37,5	62,5	0,0	0,0
24	CER	B, D, Q, W, X, Y	0,0	100	0,0	0,0
25	Calçada	<=1 (sem calçada e ruim)	0,0	100	0,0	0,0
26	Calçada	>1 (boa e bonita)	12,5	62,5	0,0	25,0
27	CER	I, K, M, P, R, AE	6,8	45,5	29,5	18,2
28	CER	O, T, AA, AF	9,1	77,3	9,1	4,5
29	CER	I, M, N, P, AA	7,1	50,0	32,1	10,7
30	CER	K, O, R, T, AF	14,3	7,1	57,1	21,4

D = desnutrição; E = eutrofia; O = obesidade; S = sobrepeso; linha dos nós terminais destacados em cinza.

Esta árvore resultou em 16 nós terminais, sendo o nó terminal 5 de 2º nível, 7 e 12 de 3º nível, os nós terminais 14, 15 e 18 de 4º nível e os nós terminais 21, 22, 23, 24, 25, 26, 27, 28, 29 e 30 de 5º nível. A cada ramificação foi calculado o valor de *aprimoramento*, que no caso desta árvore variou de 0,001 a 0,012. O *aprimoramento* mede essencialmente o quanto os nós filhos aprimoraram, em relação aos nós pais, considerando a homogeneidade segundo a variável dependente.

A partir da raiz, Nô 0, a árvore se ramifica em dois nós filhos, 1 e 2, que segregam os CERs em dois grupos. Os CERs contidos em cada nó apresentam características homogêneas quanto à ocorrência da variável dependente estudada: estado nutricional. A Figura 4 nos permite visualizar à qual nó do 1º nível da AD cada CERs pertence, a distribuição espacial dos CERs de acordo com o nó relacionado, bem como a ocorrência dos estados nutricionais nestes dois nós.



**Figura 4. Grupos de CERs de acordo com o 1º nível da AD**

Nesta Figura, é possível verificar que há predominância de eutrofia (73,1%) nos CERs A, B, D, H, L, Q, S, V, W, X e Y presentes no Nô 1. Em contrapartida, a ocorrência de eutrofia nos CERs que compõe o Nô 2 (C, E, F, G, I, J, K, M, N, O, P, R, T, U, Z, AA, AB, AC, AE e AF) limita-se a 53,2%. A Eutrofia corresponde ao estado nutricional adequado aos pré-escolares, consequentemente, no Nô 1 observa-se a ocorrência de condições nutricionais mais adequadas.

A variável que melhor explica a variabilidade dos dados segundo o estado nutricional é “CER”, correspondente à unidade escolar que o pré-escolar frequenta. Na AD (Figura 3) é possível observar que 8 das 15 bipartição envolveram esta variável (CER). A Tabela 4 relaciona as variáveis independentes de acordo com sua importância para a classificação.

De acordo com os dados da Tabela 4, depois de “CER”, as variáveis que melhor explicam a variabilidade dos dados se relacionam aos indicadores socioeconômicos “Escolaridade” e “Número de pessoas”. Quanto às variáveis relacionadas à infraestrutura, as que melhor explicam a variabilidade dos dados são “Calçada” e “Padrão da Casa”.

Observando os Nós 3 e 4, explicados pela variável “Número de pessoas”, percebe-se que nas residências com maior número de pessoas a ocorrência de desnutrição é de 15,1%, contra apenas 3,7% nas moradias com o menor número de residentes. Reforçando este padrão, obesidade e sobrepeso somam 26,0% em residências com número de pessoas  $\leq 4,5$ , enquanto somam apenas 5,7% em residências com número de pessoas  $> 4,5$ .

**Tabela 4 Relação da importância das variáveis independentes selecionadas pelo algoritmo CART**

Variável Independente	Importância	Importância Normalizada
CER	0,041	100,0%
Escolaridade	0,013	32,3%
Número de pessoas	0,010	23,9%
Calçada	0,010	23,0%
Sexo	0,009	22,5%
Padrão da Casa	0,009	21,7%
Asfalto	0,006	15,6%
Rio (confrontação)	0,004	10,4%
Recuo	0,002	4,9%
Guia	0,002	4,7%
Muro/Grade	0,002	3,7%
Coleta de Esgoto	0,001	1,9%

Analizando também a variável socioeconômica “Escolaridade”, que resultou nos Nós 9 e 10 da AD, evidencia-se a ocorrência de desnutrição entre analfabetos ou primário incompleto e primário completo ou ginásial incompleto (29,2%).

Se for considerada a variável de infraestrutura “Calçada”, que classificou os pré-escolares no Nós 13, correspondente aos locais sem calçada, pode-se constatar que neste grupo ocorre 50,0% de eutrofia. Em contrapartida, no Nós 14, correspondente aos locais com calçada, a eutrofia aumenta para 78,0%. Consequentemente, os estados nutricionais indesejados, especificamente sobre peso e obesidade, têm maior ocorrência em locais sem calçada.

Prosseguindo com interpretações similares, analisando-se os Nós 19 e 20, explicados a partir da variável “Padrão da Casa”, observa-se que sobre peso e obesidade se destacam. No Nós 19, correspondente aos pré-escolares que vivem nas casas de péssimas condições, a ocorrência de obesidade e sobre peso somam apenas 36,30%. Enquanto no Nós 20, referente aos pré-escolares que vivem em casas de melhores condições, a frequência de ocorrência de obesidade e sobre peso somam 54,8%.

#### 4 CONCLUSÕES

De forma geral, pode-se concluir que a variável que melhor explicou a variabilidade dos dados do estado nutricional corresponde ao CER que o pré-escolar frequenta. Dentre os dois nós formados a partir da primeira bipartição da árvore de decisão, as melhores condições nutricionais concentraram-se nos CERs do Nós 1. A árvore de decisão revelou também que os nós mais puros foram explicados pelas variáveis CER e Calçada.

Apesar da variável CER que o pré-escolar frequenta ter sido a que melhor explicou a variabilidade dos dados, as variáveis socioeconômicas referentes à escolaridade e número de pessoas residentes com o pré-escolar, também tiveram participação relevante na explicação dos dados, respectivamente 32,3% e 23,9% (considerando a importância normalizada). De forma geral, a ocorrência de desnutrição foi maior tanto nas residências com o maior número de moradores, quanto nas famílias com menor nível de escolaridade.

As variáveis de infraestrutura relacionadas à calçada e padrão da casa do pré-escolar explicaram respectivamente 23,0% e 21,7% da variabilidade dos dados (considerando a

importância normalizada). Tratando-se da variável correspondente à calçada, observou-se que a presença da mesma está relacionada a maiores ocorrências de estado nutricional adequado, conforme observado nas porcentagens de eutrofia nos nós resultantes da variável calçada. Já a variável correspondente ao padrão da residência do pré-escolar, explica mais claramente a questão do sobre peso e obesidade, que se apresentam com maior frequência nos pré-escolares que residem em casas de melhor padrão.

Apesar de indicadores embasados no IMC serem mais frequentemente utilizados na avaliação do estado nutricional de pré-escolares, indicadores socioeconômicos e de infraestrutura também colaboraram na compreensão da ocorrência de desnutrição, eutrofia, sobre peso e obesidade. Os resultados fornecidos pela técnica de mineração de dados (Árvore de Decisão) com base no algoritmo CART indicaram que a incorporação de indicadores alternativos relacionados à estado nutricional são promissores e podem complementar o subsídio a tomada de decisão no planejamento e gestão da saúde pública.

Com base nos resultados observados, estudos futuros poderiam considerar a inclusão de algumas destas variáveis socioeconômicas e de infraestrutura urbana como complemento ao indicador tradicional relacionado ao peso e altura. Escolaridade, número de pessoas, presença de calçada e padrão da casa se mostraram variáveis promissoras neste sentido. Outro ponto relevante é que, como a variável que melhor explicou o estado nutricional foi o CER que o pré-escolar frequenta, as políticas públicas de governança nutricional poderiam direcionar suas estratégias de acordo com o indicador nutricional de cada instituição, considerando o CER como um tipo de unidade de gestão.

## 5 AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 e com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## 6 REFERÊNCIAS

ABEP (2008) Critério de classificação econômica Brasil, **Associação Brasileira de Empresas de Pesquisa**. Disponível em: <<http://www.abep.org/criterio-brasil>>. Acesso em: 24 nov. 2019.

Barbosa Filho, V. C., Quadros, T. M. B., Souza, E. A., Gordia, A. P. e Campos, W. (2010) A utilização do critério da Organização Mundial de Saúde para classificação do estado nutricional em crianças, **Motriz**, Rio Claro, UNESP, 16(4), 811–819.

Barcellos, C. de C. e Ramalho, W. M. (2002) Situação atual do geoprocessamento e da análise de dados espaciais em saúde no Brasil, **Informática Pública**, 4(2), 221–230.

Breiman, L., Friedman, J.H., Olshen, R.A. e Stone, C.J. (1984) **Classification and Regression Trees**, Wadsworth International Group, Belmont, CA.

Druck, S., Carvalho, M.S., Câmara, G. e Monteiro, A.V.M. (2004) **Análise espacial de dados geográficos** (eds.), Embrapa, Brasília, ISBN: 85-7383-260-6.

Fajersztajn, L., Veras, M. e Saldiva, P. H. N. (2016) Como as cidades podem favorecer ou dificultar a promoção da saúde de seus moradores?, **Estudos Avançados**, 30(86), 7-27.

Foschini, A. e Campos, J. (2010) Indicadores Antropométricos do Estado Nutricional de Pré-Escolares em Araraquara-SP, **Alimentos e Nutrição**, 21(3), 349-355.

IBGE (2020) **Cidades e Estados: Araraquara**. Disponível em: <<https://www.ibge.gov.br/cidades-e-estados/sp/araraquara.html>>. Acesso em: 06 set. 2020.

Kass, G.V. (1980) An exploratory technique for investigating large quantities of categorical data, **Applied Statistics**, 29(2), 119–127.

Medeiros, A. R. C., Araújo, Y. B., Vianna, R. P. T. e Moraes, R. M. M. (2014) Modelo de suporte à decisão aplicado à identificação de indivíduos não aderentes ao tratamento anti-hipertensivo, **Saúde em Debate**, 38(100), 104–118.

Pitombo, C. S., Souza, A. D. e Lindner, A. (2017) Comparing decision tree algorithms to estimate intercity trip distribution, **Transportation Research Part C-emerging Technologies**, 77, 16-32.

PNUD, IPEA e Fundação João Pinheiro (2010) **Atlas do desenvolvimento Humano do Brasil: Araraquara**, Programa das Nações Unidas para o Desenvolvimento. Disponível em: <[http://www.atlasbrasil.org.br/2013/pt/perfil\\_m/araraquara\\_sp#idh](http://www.atlasbrasil.org.br/2013/pt/perfil_m/araraquara_sp#idh)>. Acesso em: 02 set. 2020.

Prefeitura Municipal de Araraquara (2020) **Educação Infantil - CERs**. Disponível em: <[http://www.araraquara.sp.gov.br/governo/secretarias/educacao/paginas-educacao/copy\\_of\\_educacao-infantil](http://www.araraquara.sp.gov.br/governo/secretarias/educacao/paginas-educacao/copy_of_educacao-infantil)>. Acesso em: 06 set. 2020.

Quinlan, R. (1983) Learning efficient classification procedures and their application to chess end-games. **Machine Learning: An Artificial Intelligence Approach**, Tioga, Palo Alto, 463–482.

Sellen, D. (1998) Physical Status: The Use and Interpretation of Anthropometry, **Journal of Biosocial Science**, Cambridge University Press, 30(1), 135–144.

Souza, N. C., Pitombo, C. S., Cunha, A. L., Larocca, A. P. C. e Almeida Filho, G. S. (2017) Modelo de classificação de processos erosivos lineares ao longo de ferrovias através de algoritmo de árvore de decisão e geotecnologias, **Boletim de Ciências Geodésicas**, 23, 72–86.

WHO (2014) **Indicators for the Global Monitoring Framework on Maternal, Infant and Young Child Nutrition**, Geneva, Switzerland. Disponível em: <[https://www.who.int/nutrition/topics/indicators\\_monitoringframework\\_miycn\\_background.pdf](https://www.who.int/nutrition/topics/indicators_monitoringframework_miycn_background.pdf)>. Acesso em: 24 out. 2019

WHO (2020) **Global nutrition Report**, Bristol, UK: Development Initiatives, ISBN 978-1-9164452-7-7.