



Revista Brasileira de Geografia Física

Homepage: <https://periodicos.ufpe.br/revistas/rbgfe>



Os preditores da desigualdade socioterritorial no estado de Goiás, Brasil

Nadyelle Curcino do Carmo¹, Fabrizia Gioppo Nunes², Alex Mota dos Santos³, Carlos Fabrício Assunção da Silva⁴, Pedro Monteiro de Almeida Junior⁵

¹Doutora em Geografia, Instituto de Estudos Socioambientais, Universidade Federal de Goiás, Professora do Instituto Federal de Brasília. E-mail: nadyellecarmo@gmail.com ORCID: <https://orcid.org/0000-0001-7971-9090> ² Prof. Dr. Instituto de Estudos Socioambientais, Universidade Federal de Goiás. E-mail: fabrizia.iesa.ufg@gmail.com ORCID: <https://orcid.org/0000-0001-6159-4701> ³ Prof. Dr. Centro de Formação em Ciências Agroflorestais, Universidade Federal do Sul da Bahia. E-mail: alexmota@gfe.ufsb.edu.br ORCID: <https://orcid.org/0000-0002-5156-3968> ⁴ Prof. Dr. Departamento de Engenharia Cartográfica, Universidade Federal de Pernambuco. E-mail: carlos.assuncao@ufpe.br ORCID: <https://orcid.org/0000-0001-7009-8996> ⁵ Prof. Dr. Universidade Estadual da Paraíba. E-mail: pedroalmeida@servidor.uepb.edu.br ORCID: <https://orcid.org/0000-0003-0887-1639>

Artigo recebido em 05/07/2025 e aceito em 15/12/2025

RESUMO

Este artigo analisa os preditores da desigualdade socioterritorial no estado de Goiás, Brasil, com ênfase na influência do setor primário do agronegócio sobre os níveis de desenvolvimento municipal. Para isso, foram utilizadas 37 variáveis socioeconômicas e territoriais referentes aos anos de 2021 e 2022. A metodologia adotada combina modelos estatísticos (MQO) e algoritmos de *machine learning* (*Random Forest* e *XGBoost*). O modelo de Mínimos Quadrados Ordinários (MQO) foi empregado para examinar relações estatísticas entre variáveis socioeconômicas, enquanto os algoritmos *Random Forest* e *XGBoost* foram utilizados para avaliar a importância das variáveis na predição do Índice de Desenvolvimento Municipal (IDM). As análises consideraram indicadores como infraestrutura de saúde, acesso a saneamento, níveis de alfabetização e inserção econômica. Os resultados indicam que variáveis como empregos, leitos hospitalares e saneamento básico têm forte impacto na desigualdade socioterritorial, demonstrando que a ausência de acesso a serviços essenciais está diretamente associada a baixos níveis de desenvolvimento. Além disso, a pesquisa revela que, embora o agronegócio impulse o crescimento econômico estadual, ele também contribui para a ampliação das desigualdades socioterritoriais, ao concentrar riqueza em determinados municípios. Esses resultados evidenciam a necessidade de políticas públicas que promovam um desenvolvimento mais equitativo.

Palavras-chave: Agronegócio; Índice de Desenvolvimento Municipal (IDM); Regressão MQO; *Machine learning*; Políticas públicas.

The Predictors of Socio-territorial Inequality in the State of Goiás, Brazil

ABSTRACT

This article analyzes the predictors of socio-territorial inequality in the state of Goiás, Brazil, with an emphasis on the influence of the primary agribusiness sector on municipal development levels. To this end, 37 socioeconomic and territorial variables from the years 2021 and 2022 were used. The adopted methodology combines statistical models (Ordinary Least Squares – OLS) and machine learning algorithms (Random Forest and XGBoost). The OLS model was employed to examine statistical relationships between socioeconomic variables, while the Random Forest and XGBoost algorithms were used to assess the importance of these variables in predicting the Municipal Development Index (MDI). The analysis considered indicators such as health infrastructure, access to sanitation, literacy levels, and economic inclusion. The results indicate that variables such as employment, hospital bed availability, and basic sanitation have a strong impact on socio-territorial inequality, demonstrating that a lack of access to essential services is directly associated with low development levels. Furthermore, the research reveals that although agribusiness drives economic growth at the state level, it also contributes to the deepening of socio-territorial inequalities by concentrating wealth in specific municipalities. These findings highlight the need for public policies that promote more equitable development.

Keywords: Agribusiness; Municipal Development Index; OLS regression; Machine learning; Public policy.

Introdução

A desigualdade socioterritorial é a manifestação espacial da desigualdade social, refletindo a distribuição desigual de recursos, serviços e oportunidades em diferentes áreas geográficas (Marisco, 2020). Ela é moldada por múltiplos fatores, incluindo aspectos econômicos, sociais, históricos e geográficos, entre outros. Assim, observa-se que países, regiões e indivíduos apresentam diferentes níveis de desenvolvimento. Diferenças na concentração populacional, na economia e nas atividades socioculturais contribuem para a intensificação da desigualdade entre territórios e pessoas (Rego, Costa e Alves, 2021).

Na região central do Brasil, especialmente em áreas do Bioma Cerrado, o agronegócio tem se expandido significativamente, impulsionado por vultosos investimentos econômicos e tecnologia de ponta (Santos, Guimarães e Stacciarini, 2020). Essa expansão transformou o setor em um dos principais motores da economia nacional, que se destaca pela geração de empregos e renda, além de sua relevante contribuição para a balança comercial do país (Ferreira e Lino, 2021). No entanto, conforme apontam Santos, Guimarães e Stacciarini (2020), essa forma de ocupação econômica tem acentuado desigualdades socioeconômicas e intensificado conflitos e episódios de violência no campo. Ademais, as condições impostas pelo capital financeiro no setor das *commodities* agrícolas têm agravado os impactos socioambientais, afetando profundamente a vida das comunidades tradicionais e estimulando novas formas de conflito (Jesus e Dias, 2022).

O estado de Goiás destaca-se no cenário nacional, pela alta produção agrícola (Ferreira e Guimarães, 2024). Em 2023, inclusive, onze municípios foram considerados os mais ricos do país, com base no valor da produção agrícola (Brasil, 2024). O agronegócio é um setor importante da economia, e seus segmentos primário e de prestação de serviços são os grandes protagonistas (Ramos, Satel e Lima, 2024).

Dessa forma, segundo Ramos et al. (2024a), o agronegócio pode ser compreendido por meio de uma estrutura de três níveis principais. O primeiro nível, conhecido como primário, inclui os agentes diretamente envolvidos na produção rural, tais como agricultores, pecuaristas e outros produtores rurais. O segundo nível, denominado secundário, abrange as agroindústrias e os

produtores de insumos, que desempenham funções essenciais no processamento e no fornecimento de materiais necessários para a continuidade da cadeia produtiva. E o terceiro nível, classificado como terciário, engloba as etapas de distribuição, comércio e serviços, estabelecendo a ligação entre a produção e o consumo final (Ramos et al., 2024b).

Nesta pesquisa, as análises estatísticas se concentram exclusivamente no setor primário do agronegócio, que abrange a produção de matérias-primas, incluindo o cultivo de soja, milho e cana-de-açúcar, além da criação de bovinos, aves e suínos, e a pesca e aquicultura (Ramos et al., 2024b; Ramos, Satel e Lima, 2024).

Considerando o destaque do estado de Goiás no setor do agronegócio, este artigo levanta a hipótese de que a riqueza gerada por essa atividade econômica se concentra nas mãos de uma minoria ou se é efetivamente convertida em benefícios para a população. Para o desenvolvimento da pesquisa, foram empregadas análises de preditores da desigualdade socioterritorial com foco no papel do desenvolvimento socioeconômico do agronegócio.

De modo geral, os estudos de preditores são recorrentes para a explicação de ocorrência de fenômenos ambientais, econômicos e sociais, inclusive na área da saúde (Grekousis et al., 2022; Savic et al., 2024). Apesar da ampla utilização desses estudos em diversas áreas, observa-se uma lacuna na produção científica quando se trata especificamente da desigualdade socioterritorial. As aplicações mais frequentes desses estudos concentram-se, sobretudo, na análise do desmatamento (Trigueiro, Nabout e Tessarolo, 2020; Santos et al., 2021; Guimarães et al., 2023; Silva et al., 2023a).

No caso específico do Brasil, não foram identificados estudos de preditores para a análise da desigualdade socioterritorial, que busquem relacionar as atividades agropecuárias a esse fenômeno principalmente no que diz respeito à aplicação de modelos estatísticos e algoritmos de *machine learning*.

Sobre as demais aplicações, destaca-se a pesquisa de Takakura e Massi (2022), que empregaram preditores para investigar padrões de arborização urbana relacionados a desigualdades socioterritoriais na cidade de São Paulo. Além disso, Sun et al. (2023) utilizaram essa abordagem

para analisar os efeitos heterogêneos da urbanização na desigualdade urbano-rural na China, e Martins et al. (2024) aplicaram preditores para analisar os fatores de ocorrência de incêndios no Bioma Caatinga, localizado no nordeste do Brasil.

Quanto aos aspectos metodológicos, ressalta-se o uso de métodos de estimação de variáveis para aprofundar a compreensão da desigualdade socioterritorial (Chavés-Soto et al., 2021; Martori, Lagonigro e Iglesias-Pascual, 2022). Esses modelos são particularmente relevantes por serem desenvolvidos para tratar dados que apresentam *outliers*, ou seja, valores atípicos que podem distorcer os resultados dos modelos de regressão convencionais (Fotheringham, Yang e Kang, 2017).

Nesse sentido, o Mínimos Quadrados Ordinários (MQO) é um método de estimação de variáveis amplamente utilizado para estimar a associação entre uma variável dependente e um conjunto de variáveis explicativas (Martins et al., 2024). Sua versatilidade é demonstrada em diversos estudos, como a aplicação do MQO para estabelecer relações globais entre variáveis preditivas da pobreza na Nigéria (Adeleke et al., 2023); a previsão da taxa de pobreza em Papua, Nova Guiné (Lumbantoruan, Ohyver, Moniaga, 2023); e a avaliação dos impactos do Índice de Desenvolvimento Humano e da porcentagem da população total vivendo na pobreza em Java Central, Indonésia (Mahara e Fauzan, 2021).

A análise de preditores também pode ser realizada por meio de algoritmos de *machine learning*, que têm se tornado cada vez mais relevantes em estudos da pobreza (España, 2022). De acordo com o autor, embora a pobreza seja classicamente definida e medida pela renda monetária, as técnicas de *machine learning* proporcionam abordagens inovadoras e complementares, possibilitando a classificação e

previsão da pobreza sem depender diretamente da renda como variável explicativa (España, 2022).

Estudos recentes têm explorado o uso de algoritmos de aprendizado de máquina para analisar a pobreza em diversos países, incluindo Brasil (Santolin e Oliveira, 2023), México (Rincón, 2023), Costa Rica (Solís-Salazar e Madrigal-Sanabria, 2022), Bolívia (España, 2022) e Medellín, Colômbia (Muñetón-Santa e Manrique-Ruiz, 2023). Estes trabalhos baseiam-se em dados concretos, de fontes confiáveis, como a PNAD Contínua do Brasil, e a ENIGH (*Encuesta Nacional de Ingresos y Gastos de los Hogares* - Pesquisa Nacional de Renda e Despesas dos Domicílios) do México. Coletados por meio de pesquisas domiciliares, esses dados visam aprimorar a compreensão desse fenômeno complexo e melhorar sua previsão.

Assim, considerando as possibilidades de aplicação dos métodos mencionados para o avanço dos estudos no estado de Goiás, esta pesquisa tem como objetivo analisar a desigualdade socioterritorial nos 246 municípios goianos. A análise é sustentada por um conjunto de dados e informações obtidos sobre o setor primário do agronegócio, além de dados socioeconômicos provenientes de fontes diversas. Para tanto, foram realizadas análises estatísticas, utilizando modelos de regressão e *machine learning*.

Material e métodos

Área de Estudo - abrange os 246 municípios do estado de Goiás, conforme a delimitação oficial do IBGE (2022), representada graficamente na Figura 1. De acordo com o IBGE (2025), o estado de Goiás, localizado na região Centro-Oeste do Brasil, é o sétimo maior estado do país, com uma área de 340.242,860 km². Sua população, segundo o Censo de 2022, era de 7.056.495 habitantes, o que confere ao estado o *status* de 11º mais populoso do Brasil.

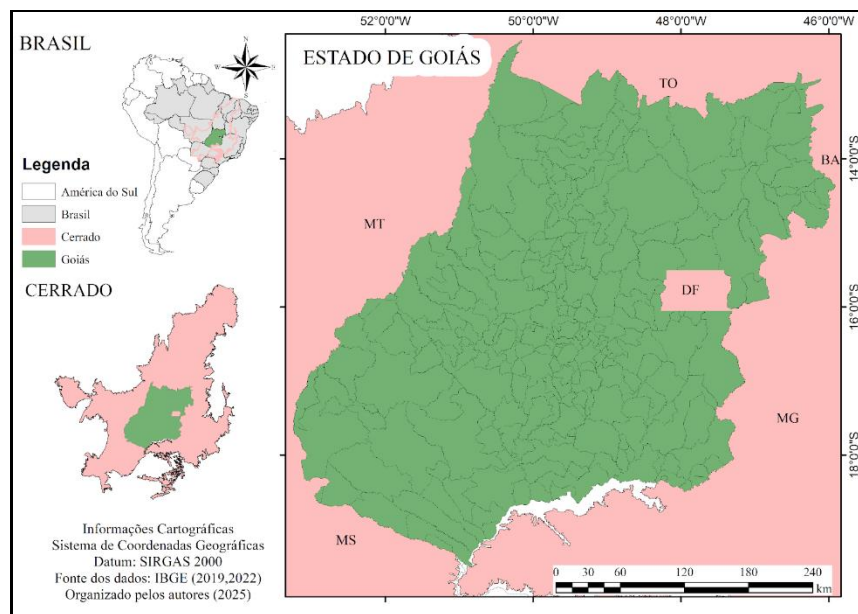


Figura 1. Localização da área de estudo.

O Índice de Desenvolvimento Humano (IDH) estadual foi de 0,737 em 2021, um valor considerado alto, já que o índice varia de zero a um (IBGE, 2025). O rendimento mensal domiciliar *per capita*, em 2024, foi de R\$ 2.098 (IBGE, 2025).

O Estado de Goiás possui a maior parte de sua área inserida no Bioma Cerrado, que cobre cerca de 70% do território goiano. Essa característica desempenha um papel essencial na configuração geográfica e socioeconômica do estado (Governo de Goiás, 2023). O clima local é tropical, com verões chuvosos e invernos secos, o que cria condições ideais para o desenvolvimento de atividades agrícolas (Governo de Goiás, 2023).

Ainda de acordo com o Governo do estado de Goiás (2023), a predominância de terras com baixa declividade, com 65% da superfície formada por chapadões e leves ondulações nas proximidades dos rios Araguaia e Tocantins, evidencia o alto potencial para o uso produtivo da terra. Essas características, alinhadas à dinâmica natural do Cerrado, criam um cenário geográfico que posiciona Goiás como um dos maiores polos de agricultura e pecuária no Brasil (Governo de Goiás, 2023).

O agronegócio é responsável por uma parcela significativa do Produto Interno Bruto (PIB) goiano, principalmente no setor de exportação de produtos agrícolas, como grãos, cana-de-açúcar e seus derivados (Resende, Guimarães e Tavares, 2025). Além disso, de acordo com os dados do Instituto Mauro Borges de

Estatísticas e Estudos Socioeconômicos de Goiás (IMB, 2024), o setor agropecuário é um dos mais relevantes para a economia goiana, visto que a produção de carne e grãos impulsionou as exportações. Desse modo, o Estado de Goiás é o quarto maior produtor nacional de grãos, com uma produção aproximada de 22,815 milhões de toneladas, o que corresponde a 9,5% da produção do país. Em 2020, os principais produtos agrícolas foram a cana-de-açúcar, a soja, o milho, o tomate e o sorgo (IMB, 2024).

Os conflitos sociais em Goiás estão intimamente ligados ao desenvolvimento econômico, à expansão do agronegócio e às transformações territoriais que ocorreram no estado ao longo do tempo (Franco e Gonçalves, 2023). As desigualdades socioeconômicas e regionais permanecem acentuadas, com a concentração de riqueza em determinadas áreas contrastando com a persistência da pobreza em outras, frequentemente desprovidas de infraestrutura adequada e políticas públicas eficientes (Queiroz e Pires, 2021).

Como instrumento principal de análise tem-se as medidas da desigualdade socioterritorial, baseadas em dados e informações agregadas ao nível de municípios, bem como dados relativos ao setor primário do agronegócio. Nessa pesquisa, foram empregados parâmetros socioeconômicos para analisar e compreender a dinâmica das condições sociais e econômicas da área de estudo, detalhados no próximo tópico.

Fontes dos dados e informações

Os dados e informações utilizados foram obtidos de fontes oficiais, ou seja, coletados, analisados e publicados por instituições públicas. As informações são compostas por variáveis que explicam a desigualdade socioterritorial, sendo a variável resposta o Índice de Desenvolvimento Municipal (IDM), conforme adaptado do estudo de Carmo, Nunes e Santos (2024).

Adicionalmente, foram incorporadas variáveis relacionadas à saúde, além de aspectos econômicos que refletem as principais atividades agropecuárias de Goiás. As variáveis explicativas utilizadas na análise estão detalhadas no Quadro 1, juntamente com suas respectivas descrições, unidades de medida e fontes.

Quadro 1: Descrição das variáveis resposta e das variáveis explicativas, com suas respectivas descrições, unidades de medida e fonte de informação.

Preditor	Descrição	Unidade	Fonte
Índice de Desenvolvimento Municipal (IDM)	Indicador utilizado para avaliar o nível de desenvolvimento de um município.	Índice	Instituto Votorantim (2022)
Área	Área dos municípios	Hectares	IBGE (2022)
Cana-de-açúcar Milho Soja Sorgo Tomate	Quantidade produzida dos principais cultivos agrícolas por município.	Toneladas	IBGE (2021)
Bois Aves	Quantidade de cabeças por município	Unidade	IBGE (2021)
Leite	Quantidade de leite produzida por município	Mil/litros	IBGE (2021)
PIB (<i>per capita</i>)	Média do valor da produção econômica por pessoa	Reais R\$ (mil)	IBGE (2021)
PIB agronegócio	Valor total de bens e serviços produzidos pelo setor agropecuário e suas cadeias relacionadas	Reais R\$ (mil)	IBGE (2021)
ICMS	Arrecadação do Imposto sobre Circulação de Mercadorias	Reais R\$ (mil)	IMB (2021)
População	Número de habitantes por município	Unidade	IBGE (2022)
Domicílios	Número de domicílios por município	Unidade	IBGE (2022)
Longevidade	Número de pessoas com 70 anos ou mais	Unidade	IBGE (2022)
Homens	Quantidade de homens por município	Unidade	IBGE (2022)
Mulheres	Quantidade de mulheres por município	Unidade	IBGE (2022)
Analfabetos	Pessoas não alfabetizadas com 15 anos ou mais	Unidade	IBGE (2022)
Homens analfabetos	Quantidade de homens não alfabetizados com 15 anos ou mais	Unidade	IBGE (2022)
Mulheres analfabetas	Quantidade de mulheres não alfabetizadas com 15 anos ou mais	Unidade	IBGE (2022)
Média Morador	Média de morador por domicílio	Unidade	IBGE (2022)
Estabelecimentos de ensino	Quantidade de estabelecimentos de ensino por município	Unidade	IMB (2021)
Número de hospitais	Quantidade de hospitais por município	Unidade	DATASUS (2022)
Leitos	Quantidade de leitos hospitalares, tanto particulares quanto geridos pelo SUS	Unidade	DATASUS (2022)
Leitos SUS	Quantidade de leitos SUS por município	Unidade	DATASUS (2022)
Leitos não SUS	Quantidade de leitos conveniados e particulares por município	Unidade	DATASUS (2022)

Leitos UTI	Quantidade de leitos de UTI por município	Unidade	DATASUS (2022)
CAD Único	Quantidade de pessoas inscritas no CADÚnico	Unidade	Ministério da Cidadania (2022)
CADÚnico Extrema Pobreza	Quantidade de pessoas em situação de extrema pobreza inscritas no CAD Único	Unidade	Ministério da Cidadania (2022)
PBF	Quantidade de pessoas inscritas no Programa Bolsa Família	Unidade	Ministério da Cidadania (2022)
Empregos	Quantidade de empregos	Unidade	RAIS (2022)
Empregos Agro	Quantidade de empregos relacionados ao setor primário do agronegócio	Unidade	RAIS (2022)
Homicídios	Quantidade de homicídios cometidos por município	Unidade	IMB(2021)
Sem água	Quantidade de domicílios sem ligação com a rede de abastecimento de água	Unidade	IBGE (2022)
Sem esgoto	Quantidade de domicílios sem ligação com a rede de coleta de esgoto	Unidade	IBGE (2022)
Sem banheiro	Quantidade de domicílios sem banheiro	Unidade	IBGE (2022)
Lixo precário	Quantidade de domicílios com coleta de lixo precária	Unidade	IBGE (2022)

Elaboração: Os autores (2025).

As variáveis econômicas consideradas neste estudo referem-se às principais culturas agrícolas e à pecuária, sendo relacionadas à área do município em km². As informações relacionadas a empregos foram obtidas a partir da Relação Anual de Informações Sociais (RAIS), gerida pelo Ministério do Trabalho e Emprego (MTE). Por outro lado, as variáveis sociais foram associadas aos dados demográficos (Quadro 1).

Os dados populacionais, a caracterização dos domicílios e as áreas dos municípios foram obtidos a partir do Censo Demográfico do Brasil, realizado pelo Instituto Brasileiro de Geografia e

Estatística (IBGE, 2022). Em seguida, todas as variáveis foram organizadas em uma planilha, utilizando como campo comum o código do município, também fornecido pelo IBGE. As variáveis foram inseridas no arquivo vetorial dos municípios por meio da função *Join* no ArcMap, versão 10.8.

Como técnica computacional foi adotado o modelo de regressão linear. Além disso, foram utilizados algoritmos de *machine learning*. Essas técnicas estatísticas serão detalhadas no próximo tópico.

Análise estatística

As análises dos dados e informações foram realizadas no software R, versão 4.4.0. Para isso, foi empregado o pacote '*stats*' (Marschner Donoghoe e Donoghoe, 2018), que possibilita o ajuste de múltiplos modelos lineares. Além disso, foram utilizados os pacotes *ggplot2*, *caret*, *lattice* e *dplyr*, também do software R, para as análises baseadas em técnicas de aprendizagem de máquina.

De modo geral, o *ggplot2* serve para a criação de gráficos estatísticos (Wickham, 2016). O *caret* é aplicado no treinamento e na avaliação de modelos de machine learning (Kuhn, 2019). O pacote *lattice* é destinado à visualização de dados por meio de gráficos condicionais (Sarkar, 2025), e

o *dplyr* é empregado na manipulação e transformação de dados (Wickham et al., 2023).

Como estratégia analítica aplicou-se o modelo de regressão linear para verificar a relação entre as variáveis, conforme os procedimentos descritos por Silva et al. (2020a) e Martins et al. (2024). O retorno desse procedimento foi a análise espacial de cada variável que explica a desigualdade socioterritorial, bem como das variáveis relacionadas às atividades agropecuárias, considerando a correlação espacial entre esses preditores e o IDM.

O modelo de regressão linear é um dos processos mais utilizados para a estimação dos coeficientes na regressão (Yadav e Prasad, 2024). Neste artigo foi aplicado com o objetivo de

descrever a relação estatística existente entre as variáveis selecionadas e o IDM. Esse modelo oferece uma perspectiva global, estimando um único conjunto de coeficientes para toda a área de estudo, o que facilita a generalização dos resultados obtidos, fornecendo insights sobre as relações gerais entre as variáveis (Hi, Mai e Shen, 2020).

Para selecionar as covariáveis do modelo proposto, foi utilizado o método de seleção *stepwise*, em combinação com o Critério de Informação de Akaike (AIC). Esse critério é um método matemático para avaliar quão bem um modelo se ajusta aos dados a partir dos quais foi gerado (Cavanaugh e Neath, 2019). Assim, foi selecionado o modelo com o menor valor de AIC, uma vez que esse método considera as estatísticas F parciais de todas as variáveis do modelo.

Adicionalmente, a escolha do modelo foi complementada pela análise do coeficiente de determinação (R^2), um indicador que expressa a proporção da variabilidade total da variável dependente explicada pelas variáveis independentes (Figueiredo Filho et al., 2011). A utilização do R^2 é justificada por sua relevância na avaliação da qualidade do ajuste dos modelos preditivos, sendo considerada uma das métricas mais difundidas para esse propósito (Gujarati, 2004).

Para um melhor ajuste, realizou-se a análise dos resíduos do modelo, visando identificar a presença de autocorrelação espacial, um fator crucial em pesquisas que trabalham com variáveis georreferenciadas. Para tal, aplicaram-se testes baseados na estatística F, que são amplamente empregados na comparação de variâncias e na detecção de padrões estruturais nos resíduos (Achten e Lessmann, 2020).

Segundo Knights e Prchkovska (2024), a regressão linear múltipla é uma técnica utilizada para explorar a relação entre uma variável dependente e múltiplos preditores independentes.

Esse método, como ferramenta de análise de dados, permite estimar a associação entre Y, a variável dependente, e X_i , um conjunto de preditores independentes ou variáveis explicativas (Martins et al., 2024).

Além disso, o MQO descreve a relação entre as variáveis, indicando se a direção da relação é positiva ou negativa (Alya et al., 2024; Figueiredo Filho et al., 2011). De acordo com esses autores, um coeficiente positivo significa que, ao manter as demais variáveis constantes, um aumento na variável independente resulta em um aumento na variável dependente; enquanto um coeficiente negativo indica que, sob as mesmas condições, um aumento na variável independente está relacionado a uma diminuição na variável dependente.

A aplicação do método *stepwise* com o AIC reiterou-se como a estratégia mais adequada para a seleção das covariáveis, possibilitando a construção de modelos estatisticamente eficientes e com elevado poder explicativo (Cavanaugh e Neath, 2019). A validação dos modelos ajustados foi consolidada pela análise do R^2 , critério que é referência na literatura para avaliação e comparação do desempenho de modelos estatísticos (Lumbantoruan, Ohlyver e Moniaga, 2023).

De acordo com Lumbantoruan, Ohlyver e Moniaga (2023), o coeficiente R^2 é utilizado como critério para avaliar e comparar o desempenho de modelos matemáticos, permitindo medir o grau de ajuste desses modelos a um determinado conjunto de dados experimentais. Quanto maior o valor de R^2 , melhor o modelo explica a variabilidade dos dados analisados.

A síntese metodológica adotada está representada na Figura 2, que apresenta de forma esquemática as etapas de preparação e análise dos dados.

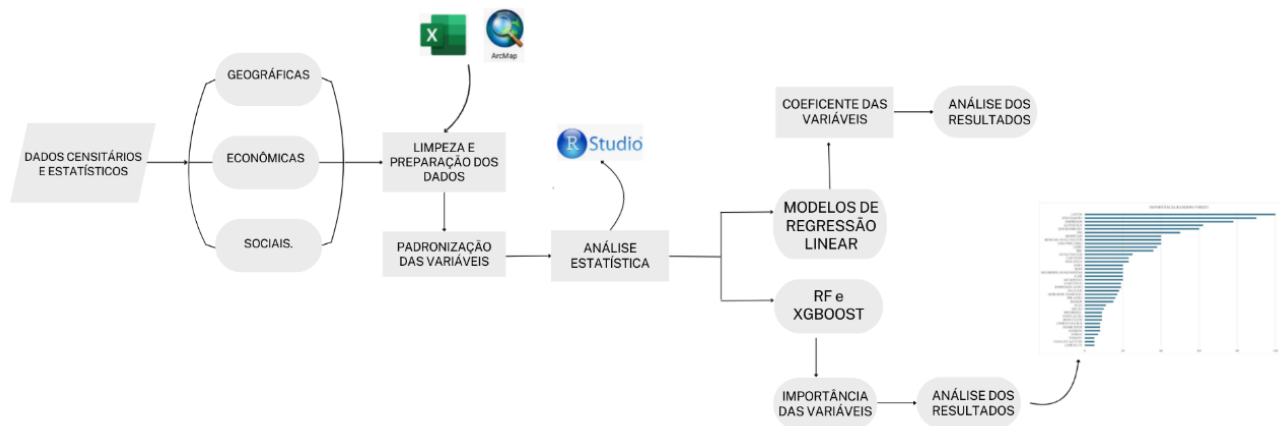


Figura 2. Fluxograma das etapas metodológicas.

Por fim, foi realizada a análise de multicolinearidade entre as variáveis independentes por meio do cálculo do Fator de Inflação da Variância (VIF), uma medida amplamente reconhecida na literatura por sua eficácia na detecção de multicolinearidade entre variáveis explicativas (Morasae et al., 2024). De modo geral, a multicolinearidade ocorre quando duas ou mais variáveis explicativas apresentam uma correlação elevada entre si (Adeleke et al., 2023). O VIF foi determinado conforme Alya et al. (2024), seguindo a equação:

$$VIF = 1 / (1 - R^2)$$

Onde:

- VIF é o Variance Inflation Factor para aquela variável.
- R^2 é o coeficiente de determinação (R-squared) de uma regressão auxiliar.

É importante salientar que a multicolinearidade refere-se a um possível problema entre as variáveis independentes ou explicativas (Alya et al., 2024). Segundo os autores, um valor de VIF superior a 10 para qualquer variável explicativa sugere um alto nível de multicolinearidade, o que pode indicar a necessidade de excluir a variável do modelo. Por outro lado, valores de VIF inferiores a 10 indicam a ausência de colinearidade entre as variáveis independentes (Alya et al., 2024).

Algoritmos de *machine learning*

Foram aplicados dois algoritmos de *machine learning* baseados em árvores de decisão, para a tarefa de classificação supervisionada, direcionada à predição da desigualdade socioterritorial: o *Random Forest* (Breiman, 2001)

e o *XGBoost* (Chen e Guestrin, 2016). A escolha desses métodos fundamenta-se na sua ampla utilização e eficácia comprovada na literatura científica para problemas de classificação complexos, especialmente em cenários que envolvem múltiplas variáveis socioeconômicas e relações não lineares (Gioia, Barros e Silva, 2022; Gioia, 2024; Shao, Ahmad e Javed, 2024).

As árvores de decisão, estrutura central desses algoritmos, operam como um conjunto de regras hierárquicas para realizar classificações ou regressões (Espana, 2022). Sua estrutura se assemelha a uma árvore, composta por nós que representam as variáveis de decisão, ramos que indicam as divisões baseadas em condições específicas, e folhas que correspondem às classes finais ou predições (Gioia, Barros e Silva, 2022).

Embora ambos os algoritmos sejam baseados na mesma estrutura conceitual, eles se diferenciam significativamente na forma como constroem e combinam as árvores (Espana, 2022). O *XGBoost* é um algoritmo que implementa o princípio do *boosting*, aprimora a precisão ao adicionar gradualmente novas árvores ao modelo construindo árvores de forma sequencial, de modo que cada nova árvore busca corrigir os erros cometidos pelas anteriores, resultando em um modelo altamente preciso (Santolin e Oliveira 2023). Essa abordagem, ao priorizar o ajuste iterativo dos erros, pode ser vantajosa em contextos em que pequenas variações nos dados podem ter impacto significativo na predição, como é o caso da análise de desigualdades socioterritoriais.

Por sua vez, o *Random Forest* adota uma abordagem diferente, a do *bagging*, construindo suas árvores de maneira independente e combinando os resultados por meio de técnicas de agregação, proporcionando maior precisão e

estabilidade ao modelo. (Espana, 2022). Esse método é conhecido por sua robustez, especialmente em cenários com dados ruidosos e multicolineares, características frequentemente presentes em bases de dados socioeconômicas (Caumo e Souza, 2021). Além disso, o *Random Forest* fornece estimativas confiáveis de importância das variáveis, auxiliando na interpretação dos principais determinantes da desigualdade socioterritorial. A aplicação combinada dos algoritmos *Random Forest* e *XGBoost* justifica-se não apenas pela sua

elevada performance preditiva, amplamente reconhecida na literatura (Imani; Beikmohammadi; Arabnia, 2025; Fatima et al., 2023), mas também pela capacidade de lidar com as especificidades dos dados socioespaciais (Gioia, 2024; Shao, Ahmad e Javed, 2024). Ademais, ambos os algoritmos demonstram flexibilidade para capturar padrões complexos e não lineares, aspectos fundamentais para a compreensão e predição das desigualdades nos territórios analisados (Fatima et al., 2023).

Resultados e discussões

As relações estatisticamente significativas entre a variável resposta e as covariáveis no modelo Mínimos Quadrados Ordinários (MQO), incluindo os coeficientes de regressão, os valores T e os níveis de significância dos coeficientes, estão apresentadas na Tabela 1. Adicionalmente, foi calculado o Fator de Inflação da Variância (VIF) para identificar a presença de multicolinearidade.

Os valores de VIF foram inferiores a 2,5, indicando ausência de multicolinearidade relevante entre os preditores.

Com base nessas observações, as covariáveis determinadas para o modelo foram as seguintes: PIB (*per capita*), Analfabetos, Leitos SUS, Leitos UTI, Empregos, Empregos Agro, Sem esgoto, Sem água, Sem banheiro e Bois. Todas com valor de $p < 0,01$.

Tabela 1: Relações estatísticas do modelo MQO para os preditores da desigualdade socioterritorial.

Variável	Estimativa	Std	Estatística T	Valor -P	VIF
(interceptar)	4,997e+00	8,650e-02	57,771	<2e-16	1,970766
PIB Agronegócio	4,032e-03	2,280e-03	1,768	0,07833	2,307556
PIB (per capita)	1,760e-06	7,119e-07	2,473	0,01412	1,594592
Analfabetos	-2,335e+00	7,994e-01	-2,921	0,00383	1,093414
Leitos SUS	7,407e+01	8,524e+00	8,689	6,57e-16	1,208338
Leitos UTI	1,966e+02	9,798e+01	2,006	0,04597	1,975957
Empregos	7,471e-01	2,491e-01	2,999	0,00300	1,456784
Empregos Agro	-7,927e-01	3,036e-01	-2,611	0,00961	1,284266
Sem esgoto	-5,320e-01	8,314e-02	-6,399	8,52e-10	1,254784
Sem água	-5,420e-01	2,228e-01	-2,433	0,01573	1,286113
Sem banheiro	-6,790e+00	2,171e+00	-3,127	0,00199	1,190907
Cana-de-açúcar	5,842e-05	3,063e-05	1,908	0,05767	1,364196
Bois	2,106e-03	4,047e-04	5,204	4,27e-07	
R ²	0,60				

*Std é o erro padrão da estimativa; VIF é o Fator de Inflação da Variância, Fonte: Dados da pesquisa.

Os coeficientes das variáveis PIB Agronegócio, PIB (*per capita*), Leitos SUS, Leito UTI, Empregos, Cana-de-açúcar e Bois

apresentaram valores positivos, indicando uma associação direta com o Índice de Desenvolvimento Municipal (IDM) nos municípios goianos. Isso significa que quanto maiores os valores dessas variáveis, maior será o

IDM. Por outro lado, as variáveis com coeficientes negativos, como Analfabetos, Empregos Agro, Sem esgoto, Sem água e Sem banheiro, demonstram uma relação inversa, ou seja, quanto maiores seus valores, menor será o IDM.

O MQO apresentou um R^2 ajustado de 0,60. Isso indica que, aproximadamente, 60% da variação do IDM pode ser explicada pelas covariáveis (variáveis explicativas) incluídas no modelo. A análise dos resíduos revelou uma autocorrelação espacial negativa (Índice Global de Moran = - 0,0381) que não foi estatisticamente significativo, e valor $p = 0,559$. Esse resultado sugere que os resíduos do modelo não estão espacialmente autocorrelacionados, o que é um bom indicativo para a robustez do modelo.

Os resultados obtidos pelo MQO evidenciam que o desenvolvimento municipal não é determinado por um fator isolado, mas sim pela interação entre variáveis econômicas, sociais e estruturais. Essa constatação corrobora com o estudo de Moraes, Swart e Jordaan (2021), que demonstraram como a diversidade de indústrias e ocupações nas estruturas produtivas regionais exerce efeitos distintos sobre a desigualdade de renda.

De forma complementar, Giomo (2023) apresenta evidências de que a complexidade econômica e o volume de empregos no setor industrial são cruciais para explicar a desigualdade de renda entre as Unidades da Federação Brasileira. Além disso, uma diminuição tanto na complexidade econômica quanto na participação relativa do emprego industrial tende a ampliar a desigualdade de renda nessas regiões (Giomo, 2023). A autora destaca ainda que para diminuir a desigualdade de renda nas Unidades da Federação Brasileira deve-se aumentar a diversidade econômica e os empregos industriais nesses locais.

Contudo, é importante observar que, em Goiás, apesar do destaque do estado na produção industrial da região Centro-Oeste, essa proeminência nem sempre se traduz uniformemente em altos índices de desenvolvimento em todos os municípios.

O analfabetismo está negativamente relacionado ao IDM, o que claramente reforça o papel fundamental da educação no desenvolvimento socioterritorial. Em geral, setores como a indústria exigem mão de obra qualificada. No agronegócio, observa-se uma tendência da necessidade de escolaridade dos trabalhadores, impulsionada pelo avanço tecnológico e pela mecanização (Ramos, 2025; Teixeira et al., 2025). Essa demanda por tecnologia tem, inclusive, gerado escassez de mão de obra qualificada no meio rural, estimulando a busca por capacitação específica nesse setor (Lima, 2021). Apesar disso, Andrade e Calencio (2024) evidenciam que municípios com economia fortemente baseada no agronegócio tendem a apresentar IDM inferior, mesmo dentro de regiões mais desenvolvidas.

Para os resultados da aplicação de *Machine Learning* foram utilizadas 246 amostras, ou seja, a quantidade de municípios goianos, para treinar e validar o modelo *Random Forest*, resultando no ranqueamento dos 37 preditores que influenciam no IDM dos municípios goianos (selecionados no Quadro 1). O gráfico da Figura 3 apresenta os preditores que mais influenciam neste índice, conforme as métricas de avaliação do *Random Forest*.

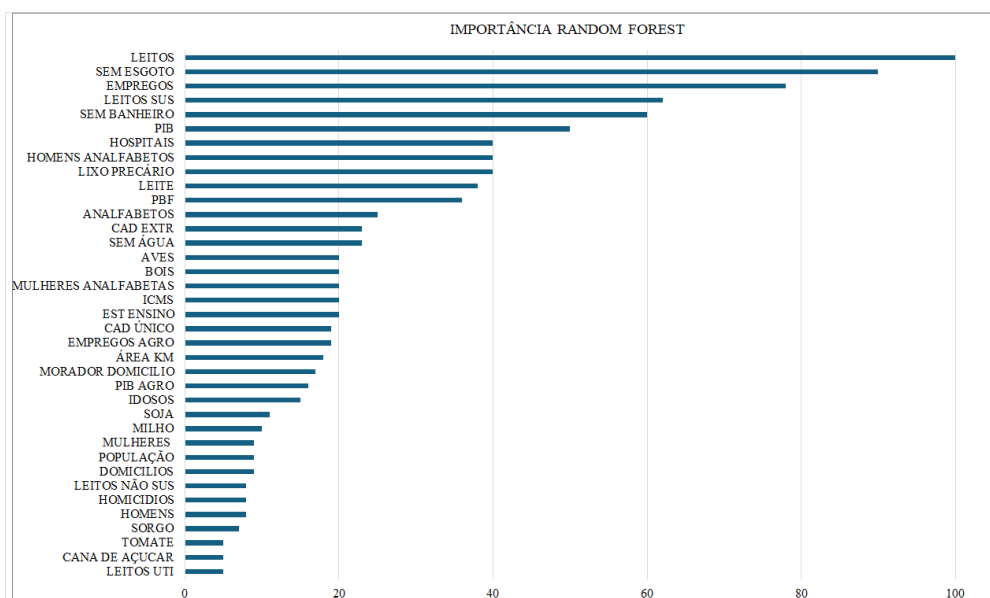


Figura 3. Importância das variáveis preditivas para o índice de desenvolvimento municipal dos municípios goianos a partir dos resultados do algoritmo *Random Forest*.

Na avaliação da importância das variáveis, os resultados indicaram Leitos, Sem Esgoto, Empregos, Leitos SUS, Sem Banheiro, PIB, Lixo precário, Homens Analfabetos, Hospitais, Quantidade de Leite Produzido, PBF, Analfabetos, Sem Água e CAD_EXTR. As demais variáveis apresentaram uma importância inferior a 20% para o modelo. As mesmas 246 amostras e os 37 preditores foram empregados para treinar e validar o modelo *XGBoost*. O objetivo foi igualmente

identificar as variáveis que mais influenciaram no IDM dos municípios goianos.

O gráfico da Figura 4 apresenta os preditores com maior influência, conforme as métricas de avaliação do *XGBoost*. As variáveis que se destacaram por sua relevância no IDM foram: Leitos, Sem esgoto, PIB, Lixo precário, Empregos, CAD_EXTR, Quantidade de Leite Produzido, Empregos Agro, Mulheres analfabetas, Leitos SUS e Cad_único. As demais variáveis apresentaram peso inferior a 20% no modelo.

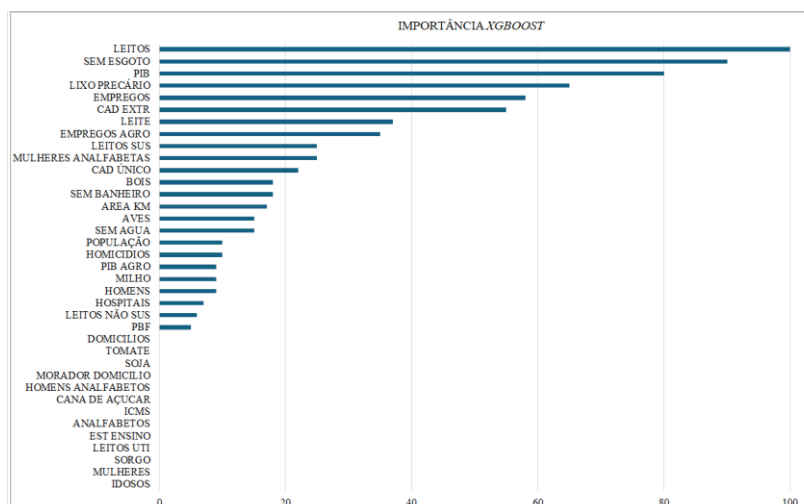


Figura 4. Importância das variáveis preditivas para o índice de desenvolvimento municipal dos municípios goianos a partir dos resultados do algoritmo *XGBoost*.

Esses resultados evidenciam que os preditores associados ao dinamismo econômico e à

infraestrutura exercem influência positiva sobre o IDM. Tal influência se justifica porque economias

mais dinâmicas e dotadas de infraestrutura adequada tendem a proporcionar melhores condições de emprego, renda e acesso a serviços públicos, promovendo, assim, maior bem-estar social e qualidade de vida. Em contrapartida, os preditores que indicam carências ou deficiências estruturais resultam em contextos que limitam o desenvolvimento humano e social, refletindo-se em menores valores de IDM.

Em ambas as avaliações, tanto *XGBoost* quanto *Random Forest*, gráfico da Figura 5 o preditor Leitos teve destaque, o que faz importante ressaltar que essa variável diz respeito aos leitos hospitalares em geral, ou seja, tanto os particulares quanto os geridos pelo SUS, conforme apontado no Quadro 1.

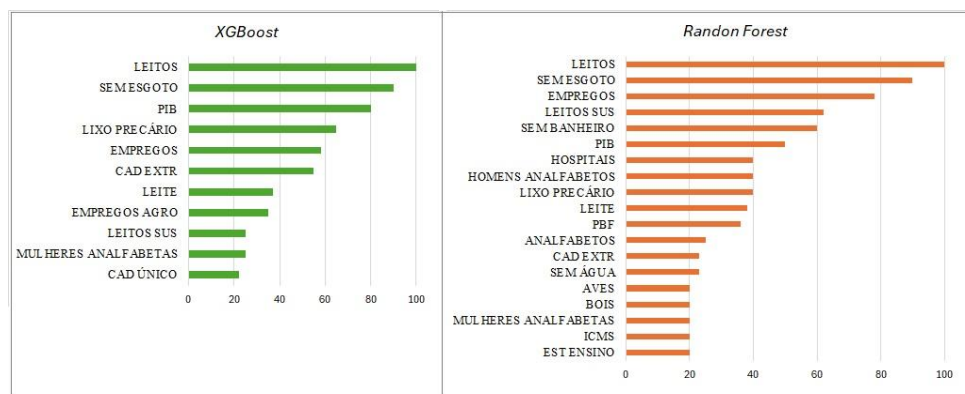


Figura 5. Comparativos dos resultados do *Random Forest* e *XGboost* para os preditores com importância acima de 20%.

O Quadro 2 apresenta uma síntese comparativa dos preditores que se destaca nas três abordagens analíticas. Dentre eles, três demonstraram uma presença recorrente em todos os métodos: Empregos, Leitos SUS e Sem esgoto. Esses podem ser considerados os principais determinantes do IDM no estado de Goiás. Além disso, outros preditores como Analfabetos,

CAD EXTR, Empregos Agro, Quantidade de Leite Produzido, Leitos, Lixo precário, PIB, Sem água e Sem banheiro aparecem em pelo menos dois dos modelos. Essa recorrência reforça a relevância dessas variáveis para a compreensão das desigualdades socioterritoriais a nível de municípios em Goiás.

Quadro 2. Preditores selecionados pelas análises.

Preditor	MQO	XGBoost	Random Forest
Analfabetos			
Bois			
CAD EXTR			
CAD ÚNICO			
Empregos			
Empregos AGRO			
Homens Analfabetos			
Hospitais			
Leite			
Leitos			
Leitos SUS			
Leitos UTI			
Lixo precário			
Mulheres Analfabetas			
PBF			
PIB			

Sem água			
Sem banheiro			
Sem esgoto			

Os resultados obtidos reforçam que o IDM é altamente sensível à infraestrutura de saúde e saneamento, assim como a fatores cruciais como a educação e a inserção econômica da população. De fato, a ausência de infraestrutura básica está diretamente associada às disparidades territoriais que observamos. Segundo Macêdo e Lima (2022), a presença de saneamento básico nos domicílios não apenas melhora a qualidade de vida, mas também reflete diretamente a renda e as condições socioeconômicas das famílias. Contudo, os autores alertam que uma parcela considerável dos domicílios goianos ainda sofre com a carência de esgotamento sanitário. Além disso, apontam que a proporção de residências com acesso a saneamento básico cresce expressivamente conforme a renda aumenta, evidenciando uma relação inequívoca entre infraestrutura sanitária e as condições socioeconômicas da população.

O analfabetismo, tanto em Goiás quanto no restante do país, reflete desigualdades estruturais profundas e falhas persistentes na garantia do

direito à educação (Souza, Fonseca e Lima, 2023). Dados da Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua) para o período de 2016 a 2023 mostram que, em 2023, cerca de 4,5% da população goiana com 15 anos ou mais ainda era analfabeta.

O mercado de trabalho no agronegócio goiano tem registrado crescimento expressivo, atingindo 1.055.201 trabalhadores em 2023 (Ramos et al., 2024). Esse número representa 26,6% da população do estado, destacando a relevância econômica do setor para Goiás. No entanto, como visto em nossos resultados, a presença e expansão do agronegócio não se traduz sozinha em melhorias nos índices de desenvolvimento humano. Esse cenário aponta para uma necessidade urgente de políticas públicas que consigam articular o desenvolvimento econômico impulsionado por este setor com o bem-estar social e ambiental, além da garantia de uma melhor distribuição de renda e acesso a serviços públicos essenciais para a população goiana.

Conclusões

A pesquisa evidenciou que a desigualdade socioterritorial em Goiás é resultado de uma complexa interação entre dinâmicas econômicas, estrutura social e distribuição desigual de serviços e oportunidades. A integração entre o método estatístico (regressão MQO) e as técnicas de aprendizado de máquina (*Random Forest* e *XGBoost*) permitiu identificar um conjunto de variáveis que impactam o Índice de Desenvolvimento Municipal (IDM), com destaque recorrente para os preditores Empregos, Leitos SUS e Sem esgoto.

Os resultados demonstraram que, embora o agronegócio seja um dos principais motores da economia goiana, seus efeitos no desenvolvimento humano não se distribuem de forma equitativa. Isso repercute na concentração de renda e infraestrutura em áreas específicas, perpetuando desigualdades históricas. Essa seletividade territorial reforça os contrastes entre centros econômicos dinâmicos e regiões marginalizadas, que são marcadas por déficits de serviços públicos essenciais, como saúde, educação e saneamento.

A convergência entre as abordagens estatísticas e computacionais utilizadas destacou variáveis que se repetem como fatores críticos nas diferentes técnicas aplicadas. Analfabetismo, empregos, precariedade de saneamento básico e baixos indicadores de saúde pública permanecem fortemente associados ao baixo IDM. Tais elementos reforçam a urgência de políticas públicas capazes de conciliar o crescimento econômico impulsionado pelo agronegócio com investimentos substanciais em infraestrutura social e educacional, a fim de reduzir as disparidades regionais e assegurar uma distribuição mais justa dos benefícios do desenvolvimento.

Adicionalmente, a aplicação de algoritmos de *machine learning* demonstrou ser uma estratégia metodológica eficaz para identificar padrões complexos e não lineares relacionados à desigualdade socioterritorial, ampliando as possibilidades analíticas em relação aos modelos tradicionais. A pesquisa contribui, assim, para o avanço do campo ao evidenciar a pertinência dessas técnicas na análise de fenômenos socioespaciais.

Destaca-se que, embora o estado de Goiás apresenta índices de desenvolvimento superiores à média nacional, os resultados indicam a necessidade de uma agenda integrada de políticas públicas voltadas à promoção da justiça social e redução das desigualdades. A continuidade de estudos que combinem métodos estatísticos tradicionais com técnicas avançadas de análise preditiva poderá aprofundar ainda mais a compreensão sobre os fatores determinantes da desigualdade socioterritorial no Brasil e em outras regiões caracterizadas por dinâmicas econômicas semelhantes.

Dessa forma, compreende-se que a desigualdade socioterritorial em Goiás não é um fenômeno isolado, mas sim o resultado de um processo histórico de produção do espaço, orientado por interesses econômicos hegemônicos. Para enfrentar essa realidade, é essencial que as políticas públicas sejam formuladas com uma abordagem territorializada, capaz de reconhecer as especificidades socioespaciais de cada município goiano. A promoção de um desenvolvimento mais justo requer ações integradas, que articulem crescimento econômico, justiça social e a reconfiguração das dinâmicas territoriais.

Agradecimentos

Agradecimentos da primeira autora ao Instituto Federal de Brasília (IFB) pela concessão de licença-capacitação.

Referências

- Achten, S., & Lessmann, C. (2020). *Spatial inequality, geography and economic activity*. *World Development*, 136, 105114. <https://www.sciencedirect.com/science/article/abs/pii/S0305750X20302412>
- Adeleke, R., Alabede, O., Joel, M., & Ashibuogwu, E. (2023). *Explorando as variações geográficas e os fatores que influenciam a pobreza na Nigéria*. *Regional Science Policy & Practice*, 15(6), 1182–1198. <https://www.sciencedirect.com/science/article/pii/S1757780223007357>
- Alya, N. A., Almaulidiyah, Q., Farouk, B. R., Rantini, D., Ramadan, A., & Othman, F. (2024). *Comparison of geographically weighted regression (GWR) and mixed geographically weighted regression (MGWR) models on the poverty levels in Central Java in 2023*. *IAENG International Journal of Applied Mathematics*, 54(12), 2746–2757. https://www.iaeng.org/IJAM/issues_v54/issue_12/IJAM_54_12_25.pdf
- Andrade, S., & Calencio, M. (Coords.). (2024). *O agro pode mais: caminhos para o desenvolvimento sustentável*. Agenda Pública. <https://agendapublica.org.br/o-agro-pode-mais/>
- Brasil. Ministério da Agricultura, Pecuária e Abastecimento. (2024). *Os 100 municípios com maior participação no valor da produção da agricultura brasileira* (Nota nº 01-2024/CGPOP/DAEP/SPA/MAPA). <https://www.gov.br/agricultura/pt-br/assuntos/noticias/mapa-divulga-os-100-municipios-mais-ricos-do-agronegocio-em-2023>
- Breiman, L. (2001). *Random forests*. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Carmo, N. C., Nunes, F. G., & Santos, A. M. (2024). *Território, sociedade e desigualdade: uma revisão de literatura*. *Revista Verde Grande: Geografia e Interdisciplinaridade*, 6(1), 31–51. <https://doi.org/10.46551/rvg26752395202413151>
- Caumo, R. B., & Souza, J. A. (2021). *A qualidade de indicadores socioeconômicos produzidos a partir de Big Data*. *Planejamento e Políticas Públicas*, (57).
- Cavanaugh, J. E., & Neath, A. A. (2019). *The Akaike information criterion: background, derivation, properties, application, interpretation, and refinements*. *Wiley Interdisciplinary Reviews: Computational Statistics*, 11(3), e1460. <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wics.1460>
- Chávez-Soto, T., Cervera-Gómez, L., Vilchis-Mata, I., & Garrocho-Rangel, C. (2021). *Socio-spatial explanation of the digital divide in the intra-metropolitan area of Toluca*. *Papeles de Población*, 27(110), 159–199. https://www.scielo.org.mx/scielo.php?pid=S140574252021000400159&script=sci_abstract&lng=en
- Chen, T., & Guestrin, C. (2016). *XGBoost: a scalable tree boosting system*. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. <https://dl.acm.org/doi/abs/10.1145/2939672.2939785>
- España, C. P. (2022). *Clasificación de la pobreza en Bolivia utilizando Random Forest y*

- XGBoost. Cuadernos de Investigación Económica Boliviana*, 5(1), 73–98. <https://www.researchgate.net/profile/Cristian-Paucara/publication/378521356>
- Fatima, S., Hussain, A., Amir, S. B., Ahmed, S. H., & Aslam, S. M. H. (2023). *XGBoost and random forest algorithms: an in depth analysis. Pakistan Journal of Scientific Research*, 3(1), 26–31. <https://pdfs.semanticscholar.org/8abb/99db16d6ccc080e3c2a2bb207c9ea20cd5e6.pdf>
- Ferreira, R. M., & Guimarães, R. S. (2024). *Política agrícola e o desenvolvimento econômico e sustentável no agronegócio do estado de Goiás. Observatório de la Economía Latinoamericana*, 22(3), e3962. <https://ojs.observatoriolatinoamericano.com/ojs/index.php/olel/article/view/3962>
- Ferreira, R. M., & Lino, E. N. S. (2021). *Expansão agrícola no Cerrado: o desenvolvimento do agronegócio no estado de Goiás entre 2000 e 2019. Caminhos de Geografia*, 22, 1–17. <https://doi.org/10.14393/RCG227951217>
- Figueiredo Filho, D., et al. (2011). *O que fazer e o que não fazer com a regressão: pressupostos e aplicações do modelo linear de mínimos quadrados ordinários (MQO). Revista Política Hoje*, 20(1), 44–99.
- Fotheringham, A. S., Yang, W., & Kang, W. (2017). *Multiscale geographically weighted regression (MGWR). Annals of the American Association of Geographers*, 107(6), 1247–1265. <https://www.tandfonline.com/doi/abs/10.1080/24694452.2017.1352480>
- Gioia, T. B. (2024). *Padrões espaço-temporais e fatores associados às doenças negligenciadas no estado de Goiás* (Tese de doutorado). Universidade Federal de Goiás.
- Grekousis, G., Feng, Z., Marakakis, I., Lu, Y., & Wang, R. (2022). *Ranking the importance of demographic, socioeconomic, and underlying health factors on US COVID-19 deaths: a geographical random forest approach. Health & Place*, 74, 102744. <https://doi.org/10.1016/j.healthplace.2022.102744>
- Guimarães, P. O., et al. (2023). *Spatial analysis of deforestation factors in the Atlantic Forest Biome/Brazil. Revista GEOgrafias*, 19(1), 1–19. <https://periodicos.ufmg.br/index.php/geografia/s/article/view/42662>
- He, X., Mai, X., & Shen, G. (2020). *Poverty and physical geographic factors: an empirical analysis of Sichuan Province using the GWR model. Sustainability*, 13(1), 100. <https://www.mdpi.com/2071-1050/13/1/100>
- IBGE – Instituto Brasileiro de Geografia e Estatística. (2023). *Censo demográfico 2022: resultados preliminares*. IBGE. <https://www.ibge.gov.br>
- IBGE – Instituto Brasileiro de Geografia e Estatística. (2023). *Painel de indicadores da PNAD contínua: 2016–2023*. IBGE. <https://painel.ibge.gov.br/pnadc/>
- IBGE – Instituto Brasileiro de Geografia e Estatística. (2025). *Panorama – Goiás*. IBGE. <https://cidades.ibge.gov.br/brasil/go/panorama>
- IMB – Instituto Mauro Borges de Estatística e Estudos Socioeconômicos. (2024). *Sobre Goiás: visão geral*. <https://goias.gov.br/imb/sobre-goias/#composicao-pib>
- Jesus, J. N., & Dias, W. A. (2022). *O Cerrado e as contradições do modelo de desenvolvimento econômico. Revista Mirante*, 15(2), 39–57.
- Knights, V., & Prchkovska, M. (2024). *From equations to predictions: understanding the mathematics and machine learning of multiple linear regression. Journal of Mathematics and Computational Applications*, 3(2), 1–8. <https://www.researchgate.net/profile/Marija-Prchkovska/publication/379564737>
- Kuhn, M. (2019). *The caret package*. <https://topepo.github.io/caret/>
- Lima, M. L. A. (2021). *Capacitação na modalidade EAD para o público rural: importante contribuição para a falta de mão de obra qualificada no campo em Goiás* (Dissertação de mestrado). Pontifícia Universidade Católica de Goiás.
- Lumbantoruan, J. F. A. K., Ohwyer, M., & Moniaga, J. V. (2023). *Developing a poverty model in Papua using geographically weighted regression. Procedia Computer Science*, 227, 243–252. <https://www.sciencedirect.com/science/article/pii/S1877050923016885>
- Macêdo, M. R., & Lima, A. F. R. (2022). *Desigualdade em Goiás: um diagnóstico em busca do desenvolvimento sustentável* (Informe técnico). Governo do Estado de Goiás; Instituto Mauro Borges. https://goias.gov.br/imb/wp-content/uploads/sites/29/2024/01/repositorio_2022_029_informes_tecnicos_desigualdade_em

- [goias_um_diagnostico_em_busca_do_desenvolvimento_sustentavel_2022.pdf](#)
- Mahara, D. O., & Fauzan, A. (2021). *Impactos do índice de desenvolvimento humano e da porcentagem da população total na pobreza usando modelos OLS e GWR em Java Central, Indonésia. Eksakta: Journal of Sciences and Data Analysis*, 142–154. <https://journal.uui.ac.id/Eksakta/article/view/20448>
- Marisco, L. M. (2020). *Revisitando autores sobre os conceitos de segregação socioespacial e exclusão social na análise da produção desigual do espaço urbano. Revista Contexto Geográfico*, 5(9), 45–56. <https://www.seer.ufal.br/index.php/contextogeografico/article/view/9998>
- Marschner, I., Donoghoe, M. W., & Donoghoe, M. M. (2018). *glm2: generalized linear models*. <https://cran.r-project.org/package=glm2>
- Martins, S. F. S., et al. (2024). *The drivers of fire in the Caatinga Biome in Brazil. Forest Ecology and Management*, 572, 122260. <https://www.sciencedirect.com/science/article/abs/pii/S0378112724005723>
- Martori, J. C., Lagonigro, R., & Iglesias-Pascual, R. (2022). *Social status and air quality in Barcelona: a socio-ecological approach. Sustainable Cities and Society*, 87, 104210. <https://www.sciencedirect.com/science/article/pii/S2210670722005182>
- Morais, M. B., Swart, J., & Jordaan, J. A. (2021). *Economic complexity and inequality: does regional productive structure affect income inequality in Brazilian states?. Sustainability*, 13(2), 1006. <https://www.mdpi.com/2071-1050/13/2/1006>
- Morasae, E. K., et al. (2024). *Social determinants of spatial inequalities in COVID-19 outcomes across England: a multiscale geographically weighted regression analysis. SSM – Population Health*, 25, 101621. <https://www.sciencedirect.com/science/article/pii/S2352827324000211>
- Muñetón-Santa, G., & Manrique-Ruiz, L. C. (2023). *Prevendo a pobreza multidimensional com algoritmos de aprendizado de máquina: uma abordagem de fonte de dados aberta utilizando dados espaciais. Ciências Sociais*, 12(5), 296. <https://www.researchgate.net/publication/370700005>
- Queiroz, A. M., & Pires, M. J. S. (2021). *Crescimento econômico e a associação espacial de clusters municipais no estado de Goiás: 2002–2016* (Texto para discussão n. 2636). Ipea. <https://repositorio.ipea.gov.br/handle/11058/10525>
- Ramos, É. B. T. (2025). *Mercado de trabalho do agronegócio de Goiás: terceiro trimestre de 2024* (Boletim 002/2025). Instituto Mauro Borges. https://goias.gov.br/imb/wp-content/uploads/sites/29/2025/01/boletim_002_2025_mercado_de_trabalho_goiano_agronegocio_terceiro_tri_2024.pdf
- Ramos, É. B. T., Satel, C. I. R., & Lima, A. F. R. (2024a). *Mercado de trabalho do agronegócio de Goiás: segundo trimestre de 2024*. Instituto Mauro Borges. https://goias.gov.br/imb/wp-content/uploads/sites/29/2024/09/Boletim_017_mercado_de_trabalho_agronegocio_goiano_segundo_trimestre_2024.pdf
- Ramos, É. B. T., Rezende, G. B., Satel, C. I. R., & Lima, A. F. R. (2024b). *Mercado de trabalho no agronegócio de Goiás: evolução de 2012 a 2023*. Instituto Mauro Borges. https://goias.gov.br/imb/wp-content/uploads/sites/29/2024/04/Boletim_005_2024_mercado_de_trabalho_do_agronegocio_em_goias.pdf
- Rêgo, A. B., Costa, E. R., & Alves, L. B. (2021). *Análise econométrica da pobreza em seu aspecto multidimensional nas cinco mesorregiões do estado de Goiás. Revista de Desenvolvimento Econômico*, 1(1). <https://revistas.unifacs.br/index.php/rde/article/view/7146/4347>
- Resende, S. C. P., Guimarães, H. R. C. R., & Tavares, P. A. (2025). *Os impactos das mudanças climáticas nas cadeias produtivas do agronegócio em Goiás. Revista Tocantinense de Geografia*, 14(32), 343–361. <https://doi.org/10.70860/rtg.v14i32.19632>
- Rincón, R. (2023). *Quarterly multidimensional poverty estimates in Mexico using machine learning algorithms. Estudios Económicos*, 38(1), 3–68. <https://ideas.repec.org/a/emx/esteco/v38y2023i1p3-68.html>
- Santolin, R., & Oliveira, P. G. (2023). *Um estudo econométrico e de machine learning sobre indivíduos que se tornaram pobres na pandemia a partir da PNAD Contínua. Multitemas*, 233–257. <https://www.interacoes.ucdb.br/multitemas/article/view/4104>

- Santos, A. M., et al. (2021). *Fatores de desmatamento na Amazônia brasileira: avaliando novos preditores espaciais*. *Journal of Environmental Management*, 294, 113020. <https://www.sciencedirect.com/science/article/abs/pii/S0301479721010823>
- Santos, J. S., Guimarães, C. R., & Stacciarini, J. H. R. (2020). *Modernização do campo no Cerrado: políticas conservadoras e suas consequências*. *Espaço em Revista*, 22(2), 147–164. <https://periodicos.ufcat.edu.br/index.php/espaco/article/view/68360>
- Savić, N., et al. (2024). *Socioeconomic inequality as a predictor of unmet health needs in the older adult population of Serbia*. *Frontiers in Public Health*, 12, 1373877. <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2024.1373877/full>
- Sarkar, D. (2025). *Lattice: multivariate data visualization with R*. <https://cran.r-project.org/web/packages/lattice/index.html>
- Shao, Z., Ahmad, M. N., & Javed, A. (2024). *Comparison of random forest and XGBoost classifiers using integrated optical and SAR features for mapping urban impervious surface*. *Remote Sensing*, 16(4), 665. <https://www.mdpi.com/2072-4292/16/4/665>
- Silva, C., Melo, S., Santos, A., Junior, P. A., Sato, S., Santiago, K., & Sá, L. (2020). *Spatial modeling for homicide rates estimation in Pernambuco State, Brazil*. *ISPRS International Journal of Geo-Information*, 9(12), 740. <https://www.mdpi.com/2220-9964/9/12/740>
- Silva, C. F. A., et al. (2023a). *Spatial modelling of deforestation-related factors in the Brazilian semi-arid biome*. *International Journal of Environmental Studies*, 80(4), 1021–1040. <https://www.tandfonline.com/doi/abs/10.1080/00207233.2022.2099109>
- Silva, C. B., Gomes, F. F. B., Santos, J. V. C., & Santos, C. S. (2023b). *Uma análise comparativa das técnicas de machine learning: regressão logística, árvores de decisão, random forest e SVM*. *Apoena*, 7, 501–511. <https://publicacoes.unijorge.com.br/apoena/article/view/183>
- Solís-Salazar, M., & Madrigal-Sanabria, J. (2022). *A machine learning proposal to predict poverty*. *Revista Tecnología en Marcha*, 35(4), 84–94. https://www.scielo.sa.cr/scielo.php?pid=S0379-39822022000400084&script=sci_arttext&tlng=en
- Souza, R. B., Fonseca, M. A. R., & Lima, D. C. B. P. (2023). *O analfabetismo funcional: reflexo nas desigualdades sociais do Brasil*. *Revista Uniaraguaia*, 201–214. https://sipe.uniaraguaia.edu.br/index.php/REVI_STAUNIARAGUAIA/article/view/1245
- Sun, B., Fang, C., Liao, X., Liu, M., Liu, Z., & Guo, X. (2023). *Revealing the heterogeneous effects of new urbanization on urban-rural inequality using geographically weighted quantile regression*. *Applied Geography*, 159, 103082. <https://doi.org/10.1016/j.apgeog.2023.103082>
- Takakura, M., & Massi, K. G. (2022). *Wealth and education influences on spatial pattern of tree planting in a tropical metropolis in Brazil*. *Environmental Management*, 69(1), 169–178.
- Teixeira, F. A., et al. (2025). *Os indicadores de qualidade do emprego (IQE) no setor canavieiro nos estados de Minas Gerais, Goiás, São Paulo e Paraná a partir do Censo Agropecuário*. *Caderno Pedagógico*, 22(5), e14508. <https://ojs.studiespublicacoes.com.br/ojs/index.php/cadped/article/view/14508>
- Trigueiro, W. R., Nabout, J. C., & Tessarolo, G. (2020). *Uncovering the spatial variability of recent deforestation drivers in the Brazilian Cerrado*. *Journal of Environmental Management*, 275, 111243. <https://www.sciencedirect.com/science/article/abs/pii/S0301479720311671>
- Wickham, H. (2016). *Data analysis*. In *ggplot2: elegant graphics for data analysis* (pp. 189–201). Springer. <https://doi.org/10.18637/jss.v077.b02>
- Wickham, H., et al. (2023). *dplyr: a grammar of data manipulation* (Version 1.1.4). <https://github.com/tidyverse/dplyr>
- Yadav, V. K., & Prasad, S. (2024). *Exponential method of estimation in sampling theory under robust quantile regression methods*. *Communications in Statistics – Theory and Methods*, 53(17), 6285–6298. <https://www.tandfonline.com/doi/abs/10.1080/03610926.2023.22>