

REVISÃO SISTEMÁTICA DA LITERATURA SOBRE MODELAGEM PREDITIVA COM ENFOQUE GOVERNAMENTAL

Renata dos Santos Rodrigues Idalgo

Mestranda em Governança e Transformação Digital da Universidade Federal do Tocantins (UFT)

Analista de Sistemas do Tribunal de Justiça do Estado de Rondônia

ORCID: <https://orcid.org/0009-0005-5197-9241>

e-mail: renata.idalgo@uft.edu.br

Luciano Tenylson Nogueira Costa

Mestrando em Governança e Transformação Digital da Universidade Federal do Tocantins (UFT)

Analista de Sistemas do Tribunal de Justiça do Estado de Rondônia

ORCID: <https://orcid.org/0009-0007-5546-2156>

e-mail: luciano.tenylson@uft.edu.br

Rafael Lima de Carvalho

Doutor em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro

(UFRJ)

Mestre em Sistemas e Computação pelo Instituto Militar de Engenharia (IME)

Graduado em Ciência da Computação pela Universidade Federal do Tocantins (UFT)

Professor adjunto da Universidade Federal do Tocantins, curso de ciência da computação

ORCID: <https://orcid.org/0000-0002-5296-8641>

e-mail: rafael.lima@uft.edu.br

Recebido em: 15/10/2025

Aprovado em: 08/02/2026

RESUMO

Esta revisão sistemática investiga o uso de modelos preditivos no contexto governamental, com foco em suas aplicações, métodos, áreas de implementação e resultados relatados. Cinco questões norteadoras de pesquisa foram abordadas: os modelos preditivos mais comumente usados, os domínios da administração pública onde são aplicados, as áreas mais impactadas, as funções desempenhadas pela IA e as lacunas na literatura atual. A busca foi realizada nas bases de dados *Scopus* e *Web of Science* (2020 a 2025), resultando em 1.020 estudos, dos quais 45 atenderam aos critérios de inclusão, exclusão e qualidade. A análise revelou o uso frequente de redes neurais, árvores de decisão, florestas aleatórias e processamento de linguagem natural, aplicados em áreas como saúde, educação, finanças, planejamento urbano e governança eletrônica. Em resumo, a aplicação de modelos preditivos no contexto governamental oferece grandes potencialidades, mas também expõe desafios significativos que exigem um fortalecimento contínuo da infraestrutura e da governança. A integração desses modelos deve ser cuidadosamente monitorada para garantir a equidade e a transparência no uso de tecnologias de IA no setor público.

Palavras-chave: Modelagem preditiva; Inteligência Artificial; Administração Pública.

SYSTEMATIC LITERATURE REVIEW ON PREDICTIVE MODELING WITH A GOVERNMENTAL FOCUS

ABSTRACT

This systematic review investigates the use of predictive models in the governmental context, focusing on their applications, methods, implementation areas, and reported results. Five guiding research questions were addressed: the most commonly used predictive models, the domains of public administration where they are applied, the most impacted areas, the functions performed by AI, and the gaps in the current literature. The search was conducted in the Scopus and Web of Science databases (2020 to 2025), resulting in 1,020 studies, of which 45 met the inclusion, exclusion, and quality criteria. The analysis revealed the frequent use of neural networks, decision trees, random forests, and natural language processing, applied in areas such as health, education, finance, urban planning, and e-governance. In summary, the application of predictive models in the governmental context offers great potential, but also exposes significant challenges that require continuous strengthening of infrastructure and governance. The integration of these models must be carefully monitored to ensure equity and transparency in the use of AI technologies in the public sector.

Keywords: Predictive modeling; Artificial Intelligence; Public Administration.

1 INTRODUÇÃO

A crescente complexidade da gestão pública, aliada à disponibilidade de grandes volumes de dados armazenados nos órgãos governamentais, tem incentivado governos ao redor do mundo a adotarem tecnologias de análise preditiva e inteligência artificial (IA) para melhorar a eficiência, a tomada de decisão e a alocação de recursos. Nesse cenário, os modelos preditivos se destacam como ferramentas capazes de antecipar demandas e comportamentos, identificar padrões e apoiar decisões e políticas públicas baseadas em evidências, além da auditoria e fiscalização de utilização dos recursos públicos. Apesar da evolução tecnológica, ainda existe uma lacuna quanto à compreensão estruturada e abrangente de como esses modelos têm sido aplicados no setor público. A literatura científica é dispersa entre diferentes áreas da administração pública (como saúde, educação, segurança e planejamento orçamentário), e frequentemente carece de classificação quanto aos tipos de modelos utilizados, seus impactos reais e atividades específicas desempenhadas por IA no contexto governamental.

Diante disso, esta revisão sistemática da literatura se justifica pela necessidade de responder de forma clara e metodologicamente rigorosa questões sobre as aplicações, impactos, benefícios, limitações e desafios do uso de modelos preditivos na gestão pública, além de elencar quais atividades ou funções a inteligência artificial está exercendo. A motivação principal reside no fato de que, embora a aplicação de modelos preditivos em governos esteja em expansão, há poucos estudos que realizam um mapeamento sistemático

dessas práticas, comprometendo a construção de conhecimento consolidado, comparável e replicável. Isso dificulta tanto o avanço da pesquisa quanto a formulação de diretrizes e políticas públicas embasadas em evidência, além de não trazer clareza à população em como esses modelos têm sido amplamente utilizados e em quais tarefas. Assim, esta revisão sistemática busca preencher essa lacuna, oferecendo um panorama crítico e atualizado sobre o cenário atual da modelagem preditiva com enfoque governamental. Os resultados esperados visam apoiar pesquisadores, gestores públicos e formuladores de políticas no entendimento do potencial e das limitações dos modelos preditivos, além de indicar caminhos para futuras investigações e contribuir para o fortalecimento da transparência na gestão pública, promovendo maior clareza perante os cidadãos.

A análise do corpus selecionado revela que o uso de inteligência artificial e modelos preditivos na administração pública tem se expandido significativamente, cobrindo múltiplas áreas de atuação governamental. Entre as mais recorrentes, destacam-se saúde, educação, gestão financeira, planejamento urbano, recursos humanos e regulamentação. Os modelos preditivos mais utilizados foram: redes neurais artificiais, aplicadas em previsão de demanda de serviços e detecção de fraudes; árvores de decisão e florestas aleatórias, empregadas na análise de benefícios sociais e gestão de riscos; processamento de linguagem natural (PLN), voltado à análise de redes sociais e manifestações da população; modelos estatísticos tradicionais, usados como apoio às funções de previsão orçamentária e alocação de recursos. Já quanto às funções, que essas técnicas foram mobilizadas tem-se: apoio ou substituição de decisões rotineiras; fornecimento de avaliações matemáticas para orientar políticas públicas; otimização de resultados e previsão de demandas; gestão de benefícios sociais e fiscalização de irregularidades; policiamento preditivo e previsão de reincidência criminal; análise de ameaças à segurança.

Esta revisão sistemática tem como objetivo identificar, analisar e sintetizar os estudos acadêmicos que abordam o uso de modelos preditivos aplicados ao contexto governamental, com foco em suas aplicações, métodos utilizados, áreas de atuação (como saúde, educação, segurança e orçamento) e resultados reportados.

As questões de pesquisa elencadas foram: Quais os tipos de modelos preditivos mais utilizados na área governamental? Em quais áreas da administração pública esses modelos preditivos estão sendo utilizados? Quais são as áreas da administração pública mais impactadas pela aplicação de modelos preditivos? Quais atividades ou funções estão sendo desempenhadas pela inteligência artificial no contexto governamental? Quais lacunas a

literatura atual apresenta?

A presente revisão sistemática da literatura foi conduzida com base nas recomendações do PRISMA 2020 (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), visando assegurar transparência, reprodutibilidade e rigor metodológico em todas as etapas do processo.

Apesar dos avanços, a revisão evidencia omissões críticas, tais como: qualidade e disponibilidade de dados, limitando a robustez dos modelos; desafios éticos e regulatórios, especialmente relacionados a vieses e impactos sociais; opacidade e falta de transparência dos algoritmos; e, restrições de infraestrutura e de capacitação institucional.

2 PROCEDIMENTO METODOLÓGICO

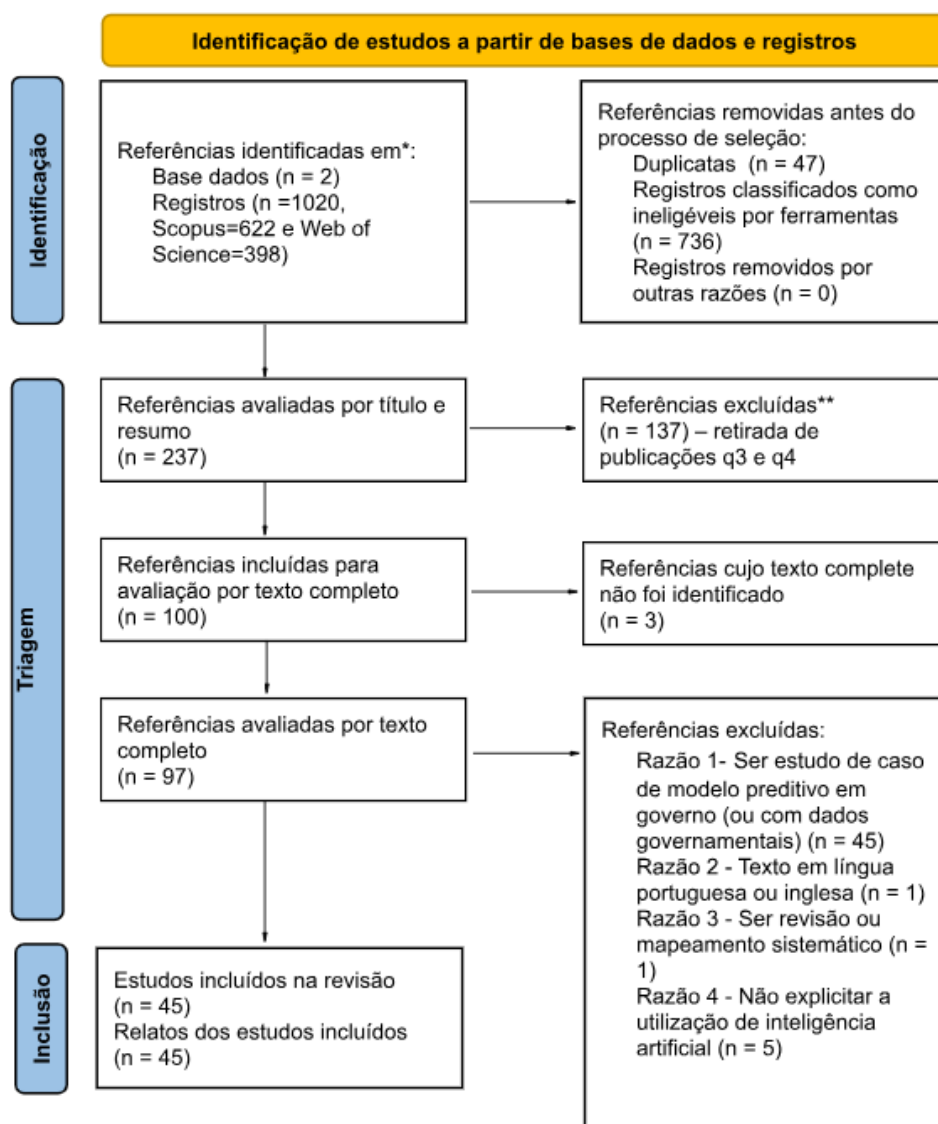
2.1 Busca e seleção de artigos

A *string* de busca definida para utilização neste estudo foi: ("predictive model" OR "predictive analytics" OR "machine learning" OR "Machine-learning" OR "random forrest" OR "Support Verctor Machine" OR "Neural Networks" OR "Gradient boost" OR "Gradient boosting" OR "KNN" OR "K-Nearest Neighbors" OR "logistic regression" OR "decision trees" OR "xgboost") AND ("public administration" OR "Government Administration" OR "Public Sector Management" OR "Public Governance" OR "Government Management" OR "Administrative Governance")

2.2 Bases de dados utilizadas

A qual foi aplicada às bases de dados *Scopus* e *Web of Science*, resultando, respectivamente, 622 e 398 artigos encontrados. A Figura 1, mostra o registro das etapas da revisão segundo o modelo PRISMA. Foram utilizados os critérios temporais (publicação ocorrida entre os anos 2020 a 2025, tendo o dia 05 de setembro de 2025 como marcador temporal para essa revisão - data da última pesquisa feita nas bases de dados) para retirada do recorte dos artigos e trabalhos revisados (*articles* e *conference papers* na base *Scopus*, e *article* e *review article* na base *Web of Science*). Além disso, foram retirados ainda os artigos que não estavam escritos em inglês ou português e não tivessem disponibilidade do texto completo já indicado nas bases de pesquisa.

Figura 1: Modelo Prisma



Fonte: Os autores (2025)

Ficaram 150 artigos na base *Scopus* e 134 da *Web of Science*, e foram marcados como duplicados 47 trabalhos pela ferramenta *Zotero* (tarefa de remoção automática dos artigos duplicados), mas foram identificados mais 15 duplicados quando da aplicação dos critérios de inclusão e exclusão, resultando em 100 estudos selecionados para aplicação desses critérios.

Após essa atividade foi realizada a remoção dos artigos que não são estudos de caso pela leitura do título e *abstract*, conforme os critérios de inclusão e exclusão.

Retirou-se 5 artigos que não explicitaram a utilização de inteligência artificial como forma da aplicação dos modelos preditivos utilizados no estudo de caso em utilização no governo ou com dados governamentais. Restando 45 artigos para análise dos critérios de qualidade.

2.3 Critérios de Seleção

Para esta revisão sistemática foram propostos 6 critérios de inclusão, 3 critérios de exclusão e 5 critérios de qualidade. Como critérios de inclusão foram elencados, nesta ordem: 1. Estar no recorte temporal de 2020 a 2025; 2. Ser estudo de caso de modelo preditivo em governo (ou com dados governamentais); 3. Ser artigo ou artigo de conferência; 4. Texto em língua portuguesa ou inglesa; 5. Texto com acesso completo; e, 6. Publicação em *Journal* Q1 ou Q2.

Já para critérios de exclusão delimitou-se: 1. Não ter o texto disponível; 2. Ser revisão ou mapeamento sistemático; 3. Não possuir revisão de pares e/ou ser literatura cinza. E ainda como critérios de qualidade elencou-se: 1. O estudo descreve explicitamente o uso de Inteligência Artificial, incluindo a aplicação de um modelo preditivo baseado em IA? (obrigatória); 2. O estudo detalha a área da administração pública de utilização da IA? (Sim - 10 pontos/Não - 0 pontos); e, 3. O estudo descreve a atividade ou função que está sendo desempenhada pela IA? (Sim - 10 pontos/Não - 0 pontos), sendo aceitos os estudos que atingiram pontuação 10 ou superior.

2.4 Avaliação da Qualidade

Para cada um dos 45 artigos que atenderam aos critérios de inclusão e exclusão, foi aplicado o formulário resultante dos critérios de qualidade, que forneceu a extração de dados dessa revisão, com as perguntas: 1. O estudo descreve o uso do modelo preditivo com IA? (obrigatória); 2. O estudo detalha a área da administração pública de utilização da IA? (Sim - 10 pontos/Não - 0 pontos); e, 3. O estudo descreve a atividade ou função que está sendo desempenhada pela IA? (Sim - 10 pontos/Não - 0 pontos).

Dados governamentais foram considerados no sentido amplo, seja que a fonte de dados fosse de fato um órgão governamental ou ainda que a população indicasse por meio das redes sociais em análise os seus anseios quanto ao governo, como no caso do artigo “Extracting Semantic Topics About Development in Africa From Social Media” (Sibitenda et al., 2024).

Foram mantidos para análise os artigos que responderam obrigatoriamente a primeira pergunta, que subsidia a resposta para a RQ1, sendo eliminados os artigos não pudessem responder qual ou quais tipos de modelos preditivos mais utilizados na área governamental

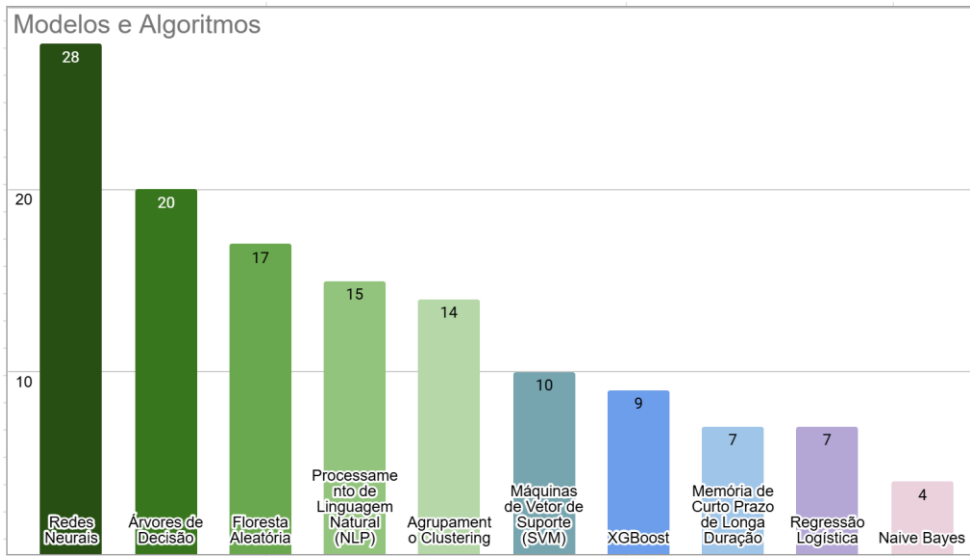
nesses últimos cinco anos. E quanto às perguntas 2 e 3, foi considerado o estudo que respondesse pelo menos uma delas.

Assim, obteve-se 40 trabalhos que responderam às três perguntas, zero que responderam somente a primeira e segunda e, por fim, cinco que responderam a primeira e a terceira pergunta.

3 SÍNTESE E ANÁLISE

Diversos estudos de caso recorreram à utilização de mais de um modelo preditivo em suas análises, como nos casos (Zaidani et al., 2025) e (Rezapour; Elmshaeuser, 2022). Assim, todas as ocorrências foram contabilizadas, adotando-se o mesmo procedimento para a contagem das áreas da administração pública em que tais modelos foram aplicados, das áreas impactadas, bem como das atividades e funções desempenhadas pela inteligência artificial.

Figura 1: Modelos e Algoritmos



Fonte: Os autores (2025)

Os dados extraídos do formulário aplicado permitiram identificar os tipos de modelos preditivos mais utilizados na área governamental, destacando-se como maior ocorrência as redes neurais e, como menor, o *naive bayes*, como ilustrado na Figura 2: Modelos e Algoritmos. A seguir, cada modelo é apresentado individualmente.

Redes neurais são ferramentas computacionais avançadas que buscam simular o processo de aprendizado e o funcionamento do sistema nervoso biológico, descritas como estruturas compostas por elementos de processamento simples e adaptativos, densamente

interconectados, chamados neurônios ou nós artificiais (Basheer; Hajmeer, 2000, p. 1). Optou-se por agrupar, em uma única categoria denominada Redes Neurais, as diferentes variações identificadas na literatura, como Redes Neurais Profundas e Redes Neurais Recorrentes, a fim de evitar uma fragmentação excessiva da classificação, como ilustrado na Figura 3. Nos estudos foram utilizadas para previsão de interações em sistemas ecologia-economia-sociedade (Peng et al., 2020, p. 1), predição de gênero e parentalidade de usuários em aplicações eletrônicas (Gatziolis; Tselikas; Moscholios, 2022, p. 4) e predição de resistência antimicrobiana (AMR) para infecções do trato urinário (Tharmakulasingam et al., 2023, p. 1).

Árvore de decisão é um algoritmo de aprendizado supervisionado utilizado para tarefas de classificação e regressão. O modelo busca particionar o espaço de atributos de forma a maximizar a pureza das folhas. Durante a inferência, uma amostra percorre a árvore até um nó folha, onde a predição é realizada (Zaidani et al., 2025, p. 11) (Orgeira-Crespo et al., 2020, p. 7) (Zhang; Zhang, 2024, p. 6). Árvores de decisão foram aplicadas para classificação de requisições em dialeto marroquino (Zaidani et al., 2025, p. 9), mapeamento do uso da Terra e identificação de parcelas agrícolas abandonadas (Morell-Monzó; Sebastiá; Estornell, 2021, p. 6) e ainda na análise e otimização de processos governamentais e estatísticas (Zhang; Zhang, 2024, p. 1).

Uma floresta aleatória (*Random Forest* - RF) é um algoritmo de aprendizado de máquina robusto e preciso, classificado como uma técnica de aprendizado em conjunto (*ensemble learning*) e um multiclassificador. É reconhecido como um dos modelos mais precisos para previsão de classificação (Orgeira-Crespo et al., 2020, p. 6) (Junfei et al., 2020, p. 6). Nos estudos de casos foram utilizadas para diversas finalidades tais como mapeamento de uso do solo (Qu et al., 2024, p. 1), avaliação de risco (Junfei et al., 2020, p. 6), (Morell-Monzó; Sebastiá; Estornell, 2021, p. 1) e (Aoki et al., 2024, p. 5), entre outros.

O Processamento de Linguagem Natural (PLN), do inglês *Natural Language Processing* (NLP), é uma tecnologia de Inteligência Artificial que tem como objetivo reduzir a lacuna entre a linguagem natural e a linguagem de máquina, utilizando o poder computacional para analisar a estrutura e a sintaxe da linguagem natural e extrair informações do conteúdo textual (Zhao; Wang; Wang, 2022, p. 5). O PLN foi utilizado para Classificação de Requisições em Dialeto Marroquino para Documentos da Administração Pública (Zaidani et al., 2025, p. 1), Melhoria de Serviços Públicos através da Análise de Feedback de Cidadãos (Modelagem de Tópicos) (Kowalski; Esteve; Mikhaylov, 2020, p. 1) e Sistemas de Busca

Jurisprudencial e Análise de Textos Jurídicos na Administração Pública Brasileira (Canedo et al., 2021, p. 4). Além disso, construiu-se uma classificação denominada agrupamento *clustering*, pois vinculou-se os algoritmos que têm em comum essa característica de ser utilizados para realizar agrupamentos: HDBSCAN, K-means, DBSCAN e *Hierarchical Clustering*.

O HDBSCAN (*Hierarchical Density-Based Spatial Clustering of Applications with Noise*) é um método de agrupamento híbrido que combina técnicas hierárquicas e baseadas em densidade. HDBSCAN foi utilizado para agrupamento de texto para identificação de tópicos em mídias sociais africanas, gerenciamento robusto de *outliers* (valores atípicos) em dados textuais e integração com o algoritmo *BERTopic* para modelagem de tópicos semânticos (Sibitenda et al., 2024, p. 7).

O *K-means* é um método de agrupamento partitivo que atribui cada ponto de dado a um dos *K-clusters*. É um algoritmo de aprendizado não supervisionado que realiza atribuições "duras" (isto é, cada ponto pertence a um único *cluster*), otimizando uma função de custo. São apresentados três estudos de caso que exemplificam a utilização do *K-means*: estimativa longitudinal de estados relacionados ao estresse a partir de dados de biossensores (Mozgovoy, 2021, p. 1), análise de dados de contratação pública para identificação de padrões e anomalias (Soylu et al., 2022, p. 15) e agrupamento de texto para identificação de tópicos em mídias sociais e documentos (Sibitenda et al., 2024, p. 2).

O DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) é um algoritmo de agrupamento cuja utilização é demonstrada para agrupar pontos de dados e identificar padrões (Sibitenda et al., 2024, p. 3). O algoritmo DBSCAN foi empregado em um estudo para comparar seu desempenho com o algoritmo *K-means* na clusterização de requisições de suporte ao cliente e agrupamento de tópicos em diversos tipos de texto (Sibitenda et al., 2024, p. 3).

O Agrupamento Hierárquico (*Hierarchical Clustering*) é um método de agrupamento realizado por meio de aprendizado não supervisionado, no qual os *clusters* (classes) são formados pela exploração das similaridades ou dissimilaridades entre os padrões de entrada, baseadas em suas intercorrelações. Esse método divide documentos em uma hierarquia, onde padrões "similares" são atribuídos ao mesmo *cluster*, conforme (Basheer; Hajmeer, 2000, p. 11). Foi utilizado para agrupamento robusto de texto para modelagem de tópicos em mídias sociais (Sibitenda et al., 2024, p. 7) e análise e classificação de padrões usando processos hierárquicos (Kowalski; Esteve; Mikhaylov, 2020, p. 5).

A Máquina de Vetores de Suporte (SVM), do inglês *Support Vector Machine*, é um modelo de aprendizado de máquina supervisionado para classificação de dados que funciona identificando o hiperplano ótimo com base no princípio do intervalo máximo no espaço de características (Xiong; Zhou, 2023, p. 6). As SVMs foram utilizadas na Classificação de Requisições em Dialeto Marroquino para Documentos da Administração Pública (Zaidani et al., 2025, p. 9), para Determinação da Similaridade de Publicações (Xiong; Zhou, 2023, p. 3) e Identificação de Terras Urbanas de Uso Misto (MUUL) a partir de Dados Geoespacial (Qu et al., 2024, p. 6).

O XGBoost (*eXtreme Gradient Boosting*) é um algoritmo de aprendizado supervisionado baseado em *ensemble* de árvores de decisão que segue o princípio do *gradient boosting* (Rezapour; Elmshaeuser, 2022, p. 7). O XGBoost foi utilizado para classificação *multilabel* de eventos acústicos urbanos (Vidaña-Vila et al., 2021, p. 15), predição precoce de evasão de alunos universitários (Borges et al., 2025, p. 9) e classificação de legislação brasileira específica da indústria (*framework* RegBR) (Valle et al., 2022, p.).

A Memória de Curto Prazo de Longa Duração (LSTM), do inglês *Long Short-Term Memory*, é uma melhoria e classe específica de redes neurais recorrentes (RNNs). Ela foi projetada para preservar dependências de longo prazo em dados sequenciais de forma mais eficiente do que as RNNs básicas. Consequentemente, as redes LSTM possuem uma enorme capacidade de armazenamento de informações e são os modelos mais comumente usados para Processamento de Linguagem Natural (PLN) (Zaidani et al., 2025, p. 12). Os estudos de caso exemplificam a utilização da Memória de Curto Prazo de Longa Duração na Classificação de Requisições em Dialeto Marroquino para Documentos da Administração Pública (Zaidani et al., 2025, p. 12), Classificação de Legislação Brasileira Específica da Indústria (*Framework* RegBR) (Valle et al., 2022, p. 18) e Predição de Evasão de Alunos Universitários (Borges et al., 2025, p. 8).

A regressão logística (LR, do inglês *Logistic Regression*) é um modelo de aprendizado de máquina supervisionado para classificação de dados (Rezapour; Elmshaeuser, 2022, p. 6) (Valle et al., 2022, p. 9). A Regressão Logística foi utilizada para Classificação Multilabel de Eventos Acústicos Urbanos (Vidaña-Vila et al., 2021, p. 15), na Predição de Resistência Antimicrobiana (AMR) em Infecções do Trato Urinário (Tharmakulasingam et al., 2023, p. 3) e Classificação de Legislação Brasileira Específica da Indústria (*Framework* RegBR) (Valle et al., 2022, p. 9).

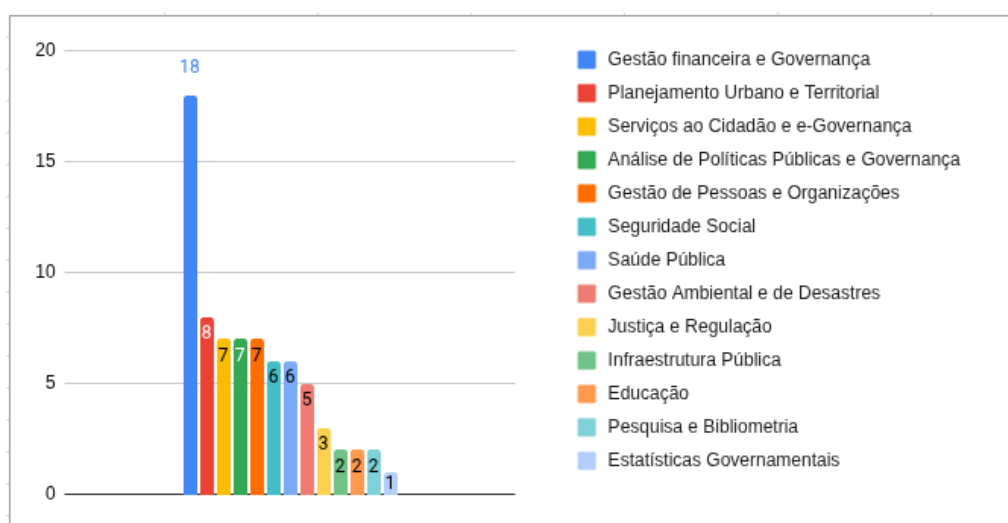
O Naïve Bayes (NB) é um algoritmo de classificação amplamente utilizado, baseado

no Teorema de Bayes, que emprega um método classificador probabilístico simples. Para exemplificar sua utilização em estudos de caso tem-se a classificação de *Tweets* em Árabe, Análise de Sentimento em *Tweets* em Dialeto Jordaniano (Zaidani et al., 2025, p. 3) e Classificação Automática de Textos de Reclamação de Serviço Ferroviário (Sibitenda et al., 2024, p. 8).

3.1 Áreas de Aplicação

Com base na revisão dos estudos selecionados, é possível observar que os modelos preditivos vêm sendo aplicados em diferentes áreas da administração pública, como gestão financeira. Descreve-se um artigo da área em que se verificou maior ocorrência de aplicações, destacando como a Inteligência Artificial tem sido utilizada. Ao final, apresenta-se um caso no contexto do Setor Público Brasileiro. A extração das áreas está consolidada na Figura 3.

Figura 3: Áreas de Aplicação



Fonte: Os autores (2025)

Na área da “Gestão financeira e Governança”, destaca-se a aplicação de técnicas de detecção de anomalias em dados de contratações públicas da Eslovênia e em registros de gastos, com o objetivo de identificar padrões suspeitos, tais como indícios de comportamento fraudulento ou de formação de monopólios. Para tanto, empregou-se o método *k-Means*, que possibilitou o agrupamento das informações e a identificação de pontos anômalos que não se ajustavam a nenhum dos *clusters* formados (Soylu et al., 2022, p. 15).

Outro estudo que merece destaque, tanto pela aplicação em pesquisa de jurisprudência

quanto pela proximidade com a realidade brasileira, é o artigo “Development and Evaluation of an Intelligence and Learning System in Jurisprudence Text Mining in the Field of Competition Defense” (Canedo et al., 2021, p. 1). A pesquisa foi desenvolvida no âmbito do Conselho Administrativo de Defesa Econômica (CADE), órgão da administração pública federal responsável pela defesa da livre concorrência.

O artigo apresenta a utilização de diferentes técnicas de Inteligência Artificial, entre elas mineração de texto, aprendizado de máquina supervisionado e não supervisionado, processamento de linguagem natural (NLP) e ontologias, aplicadas à análise de jurisprudência. Essa abordagem evidencia como os métodos de IA podem apoiar a interpretação de grandes volumes de decisões jurídicas, oferecendo subsídios relevantes para a atuação do CADE e, por consequência, para o fortalecimento da política de concorrência no Brasil.

3.2 Impactos da utilização dos modelos preditivos

Os modelos preditivos e as tecnologias associadas à inteligência artificial (IA) têm transformado diferentes esferas da sociedade, trazendo benefícios substanciais em áreas como saúde, educação e administração pública, mas também gerando desafios e repercussões negativas. Nesta seção, serão discutidos os impactos positivos e negativos observados na implementação desses sistemas, bem como uma análise de avaliações positivas que carecem de comprovação prática ou dados consolidados.

Os avanços em modelos preditivos têm gerado resultados positivos em diversas áreas, mas também é importante destacar os impactos negativos observados em alguns casos. Determinados sistemas de detecção de fraudes e de automação de decisões administrativas evidenciam falhas que resultaram em consequências graves, como erros em larga escala, discriminação e prejuízos sociais. Esses exemplos revelam os riscos associados ao uso inadequado de algoritmos e a necessidade de maior atenção às questões éticas, de transparência e de governança.

Um dos casos mais emblemáticos é o esquema australiano conhecido como *Robodebt*, que enviou aproximadamente 433.000 notificações de dívidas errôneas aos cidadãos. O sistema automatizado, que deveria identificar inconsistências em benefícios sociais, acabou por causar forte impacto social e psicológico em milhares de pessoas, sendo amplamente criticado pela imprecisão e pela ausência de mecanismos adequados de revisão humana

(Enqvist, 2024, p. 3).

Na Holanda, o sistema *SyRI* (Sistema de Informação de Risco), criado para identificar cidadãos com maior probabilidade de cometer fraudes em benefícios, foi alvo de críticas pela falta de transparência e por indícios de discriminação. Apesar de seu uso, o estudo não apresenta números de fraudes efetivamente detectadas, o que reforça a percepção de baixa efetividade e potenciais injustiças geradas por sua aplicação (Enqvist, 2024, p. 3).

Outro caso holandês de grande repercussão foi o “escândalo dos benefícios infantis” (Toeslagenaffaire), no qual cerca de 26.000 pais foram acusados erroneamente de fraude em benefícios entre 2005 e 2019. O episódio mostrou como vieses incorporados a sistemas automatizados podem levar a perseguições injustas e consequências devastadoras para as famílias, resultando em perda de confiança nas instituições públicas e em pedidos de desculpas oficiais do governo (Enqvist, 2024, p. 3).

Já na Austrália, um sistema automatizado de correspondência de dados voltado para identificar impostos e benefícios devidos levou quase meio milhão de pessoas a receberem avisos de dívida ilegais a partir de 2016. O volume de erros expôs fragilidades nos processos de verificação e destacou a necessidade de controles mais robustos depois da aplicação de modelos automatizados em larga escala (Yeung, 2022, p. 2).

Nos Estados Unidos, o estado do Arkansas implementou a automação da avaliação de benefícios por incapacidade no Medicaid, mas o processo resultou em prejuízos para cerca de metade dos beneficiários. O sistema não considerava adequadamente as condições individuais, o que levou à exclusão injusta de milhares de pessoas vulneráveis, levantando críticas severas à forma como a tecnologia foi aplicada sem salvaguardas suficientes (Yeung, 2022, p. 2).

Também nos EUA, em aplicações de justiça criminal na Flórida, um *software* utilizado para prever reincidência criminal apresentou sérios problemas. O modelo acertou apenas 20% das ocorrências, mas teve o dobro da probabilidade de prever falsamente réus negros como futuros criminosos em comparação a réus brancos, evidenciando vieses raciais e levantando um debate ético intenso sobre o uso de algoritmos preditivos em processos judiciais (Gerrits, 2021, p. 2).

Sobre os aspectos positivos, nota-se que a aplicação de modelos preditivos e tecnologias de inteligência artificial (IA) tem impactado diversas áreas da administração pública, resultando tanto em ganhos de eficiência e capacidade analítica quanto em desafios relacionados a vieses e precisão.

As principais áreas impactadas, com números e percentuais concretos, são: ° Na Suécia, um sistema baseado em regras permite que a Agência Sueca de Seguro Social emita entre 65% e 71% dos casos de benefícios parentais de forma totalmente automática. (Enqvist, 2024, p. 5).

Nos serviços públicos locais e na governança, observa-se a incorporação gradual de tecnologias preditivas e de análise automatizada de dados, como demonstrado em pesquisa realizada no Reino Unido: 27% das autoridades locais relataram estar experimentando algum tipo de análise automática de texto ou conteúdo, enquanto 17% declararam estar testando técnicas de análise preditiva, indicando uma tendência crescente de utilização dessas ferramentas para aprimorar processos administrativos e apoiar a tomada de decisão (Vogl et al., 2020, p. 4).

Na saúde pública, os modelos de *Machine Learning* (ML) vêm sendo aplicados para prever a Resistência Antimicrobiana (AMR) a partir de dados clínicos, contribuindo diretamente para a gestão de antibióticos e possibilitando intervenções mais assertivas no tratamento (Tharmakulasingam et al., 2023, p.1). Um dos impactos mais relevantes está na redução de prescrições inadequadas, já que estudos anteriores que utilizaram técnicas tradicionais de ML, como regressão logística e árvores de decisão *gradient-boosted*, em casos de infecção do trato urinário (ITU) não complicada, registraram uma diminuição de 30% na seleção de terapias antibióticas incorretas, evidenciando o potencial dessas abordagens para aprimorar tanto a eficiência clínica quanto a segurança do paciente (Tharmakulasingam et al., 2023, p. 3).

Além disso, o uso combinado de *Machine Learning* e *Deep Learning* tem se mostrado promissor em diagnósticos e no monitoramento de condições de saúde, ampliando também a capacidade de planejamento em setores específicos. Um exemplo é a detecção de Retinopatia Diabética (DR), em que um modelo baseado em *Deep Learning* (TaylorAVO-SqueezeNet) obteve desempenho expressivo, alcançando acurácia de 94,2%, sensibilidade de 94,8% e especificidade de 93,4%, resultados que evidenciam o potencial da tecnologia para apoiar diagnósticos precoces e mais precisos, reduzindo riscos de complicações em larga escala (Tharmakulasingam et al., 2023, p. 1–17).

Monitoramento de Estresse: Resultados (*Findings*). Os autores confirmam a validade convergente do *framework* conceitual proposto IW. A análise de dados empíricos sugere que modelos de dois *clusters* alcançam uma acurácia de validação cruzada quádrupla superior a 70% na identificação do estresse. O coeficiente de acurácia diminui para modelos de três

clusters, atingindo cerca de 45%. Os autores concluem que os modelos de identificação podem servir para derivar medidas longitudinais relacionadas ao estresse (Geng et al., 2022, p. 1).

Na área de gestão territorial, os modelos preditivos têm contribuído para aumentar a precisão de dados de uso do solo, impactando diretamente o cálculo de rendimentos agrícolas e apoiando políticas públicas. Um caso ilustrativo é o mapeamento de terras abandonadas, em que a aplicação do algoritmo *Random Forest* (RF) para a detecção de parcelas de citros desativadas alcançou alta precisão, registrando acurácia geral de 95% quando considerada por parcela e 87% por pixel, demonstrando o potencial da técnica para otimizar o monitoramento e a gestão de áreas produtivas (Morell-Monzó; Sebastiá; Estornell, 2021, p. 1–3).

Ainda na gestão territorial, destaca-se o impacto dos modelos preditivos na correção de inconsistências presentes em bases de dados oficiais, evitando distorções nos cálculos de rendimento agrícola. Um exemplo relevante foi a detecção de que 31% das parcelas classificadas como citros no Sistema de Informação Geográfica de Parcelas Agrícolas (SIGPAC) encontravam-se, na realidade, abandonadas, revelando como o uso de ML pode corrigir sobrestimações e fornecer informações mais fidedignas para a tomada de decisão das administrações públicas (Morell-Monzó; Sebastiá; Estornell, 2021, p. 11).

Na área de políticas ambientais, os modelos preditivos têm sido aplicados em sistemas de classificação de eventos acústicos, permitindo às administrações públicas identificar e quantificar a poluição sonora. Embora ainda não implantado, um desses sistemas foi testado e validado em ambiente real, comprovando sua viabilidade técnica ao apresentar desempenho em tempo real, com resultados de classificação gerados em menos de 1 segundo em *hardware* de baixo custo, como o Raspberry Pi Model 4B, evidenciando seu potencial para apoiar estratégias de monitoramento ambiental (Vidaña-Vila et al., 2021, p. 17).

No campo das licitações, destaca-se o projeto VigIA, desenvolvido para apoiar a Veeduría Distrital de Bogotá, órgão responsável por zelar pela moralidade e eficiência administrativa. A iniciativa trouxe impactos significativos ao otimizar a fiscalização da contratação pública e aprimorar a capacidade preditiva na detecção de irregularidades e ineficiências, enfrentando o desafio de analisar um elevado volume de contratos com recursos limitados. Entre 2018 e 2020, foram processados 87.387 contratos, considerando 66 variáveis distintas, o que representou um suporte essencial para a atuação de um pequeno órgão fiscalizador (Salazar; Pérez; Gallego, 2024, p. 6-7).

3.3 Áreas da administração pública impactadas

No presente estudo, observa-se que as áreas da administração pública mais impactadas pela aplicação de modelos preditivos correspondem diretamente àquelas em que tais modelos foram efetivamente implementados. Dessa forma, serão apresentados alguns exemplos reportados nos estudos de caso.

A utilização desses sistemas na administração de benefícios de Seguridade Social tem um impacto direto na legalidade e na distribuição justa dos benefícios, podendo resultar em efeitos negativos em larga escala devido a falhas ou vieses, como o esquema **Robodebt** na Austrália ou o sistema **SyRI** na Holanda, que gerou notificações erradas ou violou direitos de privacidade (Enqvist, 2024, p. 1).

E ainda, a utilização na Gestão Financeira Pública e Combate à Fraude impacta a otimização do planejamento orçamentário, a alocação estratégica de fundos públicos e a redução da corrupção, tornando a detecção de fraudes mais eficiente (Aldemir; Uysal, 2025).

Na dimensão de Melhoria e Personalização de Serviços Públicos e Interação com Cidadãos, os impactos observados incluem o aprimoramento da compreensão da satisfação dos usuários, a integração de suas opiniões nos processos decisórios relacionados aos serviços e a personalização das ofertas públicas, com a adaptação do conteúdo às necessidades individuais dos cidadãos (Kowalski; Esteve; Mikhaylov, 2020, p. 1).

Na área de Justiça e Jurisprudência, observa-se o impacto na promoção de maior estabilidade, uniformidade e previsibilidade nas análises jurídicas, contribuindo para o suporte às atividades dos profissionais do direito (Canedo et al., 2021, p. 1).

Em suma, as fontes tratam a utilização de modelos preditivos como a causa direta dos impactos observados nas áreas da administração pública. Portanto, as áreas de utilização e as áreas impactadas são as mesmas e se referem aos mesmos domínios de atuação do setor público. O que causa estranheza é que, na maioria dos casos, esses impactos não foram efetivamente mensurados ou quantificados, apenas relatados nos estudos de casos.

3.4 Atividades ou Funções Desempenhadas

Quanto às atividades ou funções, estas estão sendo desempenhadas pela inteligência artificial no contexto governamental são:

Tabela 1: Atividades ou Funções Desempenhadas

Categoria	Atividades / Funções
I. Governança, Gestão, Otimização e Suporte Decisório	• Mapeamento urbano (ruas, iluminação) e simulação de gêmeos digitais para políticas; • Identificação do uso da terra urbana mista (MUUL); • Monitoramento ambiental (ruído, poluição) via IoT; • Classificação e análise linguística de textos regulatórios; • Apoio a gestores: automação de processos, ETL, produtividade e eficiência em serviços; • Planejamento da força de trabalho com identificação de ocupações suscetíveis à automação; • Gestão da eficiência energética em edifícios públicos; • Identificação e previsão de surtos epidêmicos/pandêmicos com aprendizado de máquina; • Suporte a decisões baseadas em dados; • Assistência ou substituição de administradores em decisões rotineiras; • Avaliações matemáticas e otimização de resultados.
II. Segurança Pública, Justiça e Detecção de Riscos relacionados a esses temas	• Policiamento preditivo; • Previsão de reincidência criminal (ex.: COMPAS – EUA); • Previsão de ataques terroristas (ex.: SAPE – China); • Previsão de crimes em geral; • Triagem e decisões em processos de imigração.
III. Serviços Públicos e Interação com o Cidadão	• Previsão de fraudes fiscais; • Uso de agentes autônomos/chatbots para atendimento (permissões, assistência social); • Aconselhamento automatizado e transformação da comunicação cidadão-governo; • Predição de desempenho acadêmico e risco de evasão; • Avaliação da eficiência de programas de tutoria; • Análise de rotatividade de servidores na educação via folha de pagamento; • Medição do interesse público em leis; • Extração de conhecimento de mídias sociais para apoiar políticas.

Fonte: Os autores (2025)

Resumidamente, quanto às atividades ou funções é empregada para aumentar a eficiência, agilidade e racionalidade instrumental na administração pública, abrangendo desde a interação direta com o cidadão até processos internos complexos e análise de políticas. Contudo, as fontes também destacam que essa ampla utilização levanta questões éticas e legais significativas relacionadas à justiça, devido processo legal, transparência e autonomia individual (Frost, 2024).

3.5 Lacunas Identificadas

As lacunas identificadas reforçam que sua implementação efetiva e responsável exige

investimentos contínuos em infraestrutura, governança de dados, capacitação profissional e desenvolvimento de marcos normativos adequados.

Tabela 2: Lacunas Identificadas

Lacuna Identificada	Possíveis Motivos	Impactos na Administração Pública
Condições legais para sistemas automatizados	Complexidade regulatória; evolução tecnológica mais rápida que a legislação; falta de normas claras.	Insegurança jurídica e barreiras para adoção de sistemas de IA em políticas sociais.
Satisfação do usuário e feedback do cidadão	Dificuldade em harmonizar metodologias; altos custos de pesquisas longitudinais; limitações técnicas de NLP.	Falta de métricas confiáveis sobre satisfação do usuário; dificuldade de avaliar efetividade de serviços públicos.
Dados irrealistas e anotação limitada	Dificuldade logística e financeira na coleta de dados reais; escassez de iniciativas colaborativas abertas.	Baixa aplicabilidade prática dos modelos; risco de resultados enviesados.
Desequilíbrio de classes e escassez de dados estatísticos	Limitações de infraestrutura; falta de priorização política na coleta e padronização de dados.	Modelos tendenciosos; baixa confiabilidade de previsões em áreas críticas.
Exigências técnicas e computacionais para LLMs	Infraestrutura tecnológica concentrada em grandes corporações; restrições orçamentárias de governos locais.	Dificuldade de implementação de soluções avançadas de IA; aumento da dependência tecnológica externa.
Capacidade local de desenvolvimento tecnológico	Carência de equipes técnicas especializadas; predominância de terceirização; dependência de fornecedores.	Redução da autonomia governamental; soluções pouco adaptadas ao contexto local.
Conformidade com padrões de segurança (ISO/IEC 27001)	Baixa transparência das administrações; ausência de mecanismos de monitoramento e auditoria independentes.	Fragilidade na segurança da informação; risco de falhas em projetos de Cidades Inteligentes.

Fonte: Os autores (2025)

Quanto à qualidade e disponibilidade dos dados, os pontos de fragilidade encontram-se nas condições legais para sistemas automatizados na administração pública, satisfação do usuário e *feedback* do cidadão, dados irrealistas e anotação limitada e desequilíbrio de classes e escassez de dados estatísticos.

Conforme (Enqvist, 2024), observa-se ausência de estudos que aprofundem as condições jurídicas para o uso de tecnologias em processos administrativos, especialmente na concessão de benefícios de seguridade social. Além disso, princípios fundamentais, como a imparcialidade, ainda carecem de interpretações claras no âmbito do GDPR e da jurisprudência do Tribunal de Justiça Europeu (TJEU), o que dificulta a definição de diretrizes objetivas. A inexistência de legislação específica que autorize o uso de dados

sensíveis para auditoria de sistemas de IA também se apresenta como barreira, somada à indefinição do termo “estritamente necessário” no Artigo 10(5) da AIA¹. Essas lacunas permanecem, em grande parte, devido à complexidade regulatória europeia e à velocidade de evolução tecnológica, que supera a capacidade de resposta do arcabouço jurídico.

A literatura revisada mostra que a maioria dos estudos adota fatores predefinidos para explicar a satisfação do usuário, mas não valida suficientemente tais suposições. A multiplicidade de teorias utilizadas compromete a construção de um quadro robusto e comparável. Além disso, o uso de pesquisas pontuais, de alto custo e baixa taxa de resposta limita o monitoramento contínuo. Métodos como análises de opinião e modelos de tópicos, embora promissores, enfrentam vieses metodológicos e incertezas quanto à representatividade. Essas lacunas persistem principalmente devido à dificuldade em harmonizar metodologias, ao custo elevado de pesquisas longitudinais e às limitações de técnicas de processamento de linguagem natural no tratamento de dados complexos.

Muitos estudos baseiam-se em conjuntos de dados artificiais ou simplificados, construídos a partir de bibliotecas específicas, em vez de capturar fenômenos do mundo real. Isso reduz a aplicabilidade prática dos modelos. Além disso, a anotação de rótulos únicos ignora a simultaneidade de eventos, o que compromete a precisão preditiva. Esse problema decorre tanto da dificuldade logística e financeira de coletar dados reais em larga escala quanto da escassez de esforços colaborativos para criação de bases abertas e diversificadas.

A desproporção entre classes de dados (como a sub-representação de certos eventos em monitoramentos urbanos) gera vieses e compromete o desempenho dos modelos em categorias menos frequentes (Aoki et al., 2024, p. 1). A insuficiência e a baixa qualidade dos dados estatísticos governamentais agravam esse cenário. Essas lacunas refletem tanto limitações de infraestrutura de coleta e padronização quanto uma falta de priorização política na disponibilização de dados públicos consistentes e acessíveis.

Dessa forma, a ausência de dados de qualidade, o desequilíbrio entre classes e a falta de padronização metodológica dificultam avanços consistentes. Essas lacunas persistem, sobretudo, pela combinação de desafios regulatórios, limitações técnicas, custos de coleta de dados reais e barreiras políticas na governança da informação.

Já quanto aos recursos e implementação, os pontos de fragilidade encontram-se nas exigências técnicas e computacionais, capacidade local de desenvolvimento tecnológico e conformidade com padrões de segurança.

¹ AIA é a Lei da Inteligência Artificial da União Europeia, marco regulatório que disciplina a implementação de sistemas automatizados também na administração pública.

A adoção de *Large Language Models* (LLMs), como o Llama2, demanda recursos computacionais substanciais, incluindo GPUs de alto desempenho, o que limita sua implementação em administrações públicas e organizações de menor porte (Sibitenda et al., 2024, p. 14). Essa barreira é mantida pela concentração de infraestrutura tecnológica em grandes corporações e pela restrição orçamentária de governos locais.

Ainda são escassos os estudos que abordam a habilidade de governos locais em desenvolver soluções inteligentes adaptadas a seus contextos específicos, sem depender exclusivamente de fornecedores privados (Vogl et al., 2020, p. 1). Esse vazio se explica pela carência de equipes técnicas especializadas e pela predominância de contratos terceirizados, que limitam a autonomia das administrações públicas.

Há pouca evidência documentada sobre a implementação efetiva de padrões internacionais, como a ISO/IEC 27001, em projetos de Cidades Inteligentes e Gêmeos Digitais (Ciancarini; Giancarlo; Grimaudo, 2024). A escassez de estudos nesse campo pode estar relacionada à baixa transparência das administrações sobre processos internos e à ausência de mecanismos de monitoramento e auditoria independentes.

Os desafios de recursos computacionais, dependência de fornecedores externos e insuficiência de evidências sobre a adoção de padrões de segurança revelam fragilidades estruturais. Essas lacunas persistem, sobretudo, em razão de restrições financeiras, déficit de capacidade técnica interna e limitada cultura de transparência e prestação de contas.

4 CONSIDERAÇÕES FINAIS

A revisão sistemática realizada evidencia que a aplicação de técnicas de inteligência artificial na administração pública apresenta avanços significativos em diferentes áreas, como saúde, educação, gestão financeira, planejamento urbano, recursos humanos e regulamentação. Os principais modelos preditivos identificados foram redes neurais, árvores de decisão, florestas aleatórias e processamento de linguagem natural, utilizados em funções que incluem apoio e substituição de decisões rotineiras, fornecimento de avaliações matemáticas para orientar políticas, otimização de resultados, gestão de benefícios sociais, detecção de fraudes, policiamento preditivo, previsão de reincidência criminal e análise de ameaças à segurança.

Apesar dos progressos, persistem limitações importantes relacionadas à disponibilidade e qualidade dos dados, aos desafios regulatórios e éticos, à transparência dos

modelos e às restrições de recursos técnicos e institucionais, para evitar casos como cobranças indevidas do Syri e violação de privacidade como o caso holandês Robodebt. Essas lacunas reforçam que a implementação efetiva e responsável da inteligência artificial no setor público depende de investimentos contínuos em infraestrutura, governança de dados, capacitação profissional e marcos normativos adequados.

Do ponto de vista prático, este estudo contribui ao oferecer aos gestores públicos um panorama consolidado das áreas mais impactadas e das funções desempenhadas pela inteligência artificial, permitindo identificar tanto oportunidades, principalmente, de ouvir a população diretamente, como no caso africano utilizando a análise de redes sociais para extrair anseios quanto ao governo, quanto riscos associados ao uso de modelos preditivos. Do ponto de vista acadêmico, reforça-se a necessidade de pesquisas futuras que não apenas avancem na inovação tecnológica, mas também incorporem dimensões críticas de legitimidade, equidade, transparência e sustentabilidade das soluções adotadas. Assim, a inteligência artificial aplicada à gestão pública configura-se como um campo fértil para investigações interdisciplinares que dialoguem com as demandas sociais e políticas contemporâneas.

REFERÊNCIAS

- ADAMCZYK, W. B.; MONASTERIO, L.; FOCHEZATTO, A. Automation in the future of public sector employment: the case of Brazilian Federal Government. **Technology in Society**, New York, v. 67, e101722, Nov. 2021. DOI: <https://doi.org/10.1016/j.techsoc.2021.101722>.
- AHMADZAI, M.; NGUYEN, G. Differential Private Federated Learning in Geographically Distributed Public Administration Processes. **Future Internet**, Basel, v. 16, n. 7, e220, 2024. DOI: <https://doi.org/10.3390/fi16070220>.
- ALDEMIR, C.; UYSAL, T. U. Artificial intelligence for financial accountability and governance in the public sector: strategic opportunities and challenges. **Administrative Sciences**, Basel, v. 15, n. 2, e58, 2025. DOI: <https://doi.org/10.3390/admsci15020058>.
- AOKI, N.; TATSUMI, T.; NARUSE, G.; MAEDA, K. Explainable AI for government: does the type of explanation matter to the accuracy, fairness, and trustworthiness of an algorithmic decision as perceived by those who are affected? **Government Information Quarterly**, New York, v. 41, n. 4, e101980, Dec. 2024. DOI: <https://doi.org/10.1016/j.giq.2024.101980>.
- BALLESTAR, M. T.; MIR, M. C.; PEDRERA, L. M. D.; SAINZ, J. Effectiveness of tutoring at school: a machine learning evaluation. **Technological Forecasting and Social Change**, New York, v. 199, e123043, 2024. DOI: <https://doi.org/10.1016/j.techfore.2023.123043>.
- BASHEER, I. A.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. **Journal of Microbiological Methods**, Amsterdam, v. 43, n. 1, p. 3-31, Dec. 2000. DOI: [https://doi.org/10.1016/S0167-7012\(00\)00201-3](https://doi.org/10.1016/S0167-7012(00)00201-3).

BORGES, G. A.; PEDRO, C. F. S.; ANJOS, J. C.; RODRIGUES, A.; BOAVIDA, F.; SÁ SILVA, J. A platform for early class dropout prediction of university students. **IEEE Access**, Piscataway, v. 13, p. 109116-109133, 2025. DOI: <https://doi.org/10.1109/ACCESS.2025.3581751>.

CANEDO, E. D. et al. Development and evaluation of an intelligence and learning system in jurisprudence text mining in the field of competition defense. **Applied Sciences**, Basel, v. 11, n. 23, e11365, 2021. DOI: <https://doi.org/10.3390/app112311365>.

CIANCARINI, P.; GIANCARLO, R.; GRIMAUDO, G. Digital transformation in the public administrations: a guided tour for computer scientists. **IEEE Access**, Piscataway, v. 12, p. 22841-22865, 2024. DOI: <https://doi.org/10.1109/ACCESS.2024.3360408>.

CHEN, J.; LI, Q.; WANG, H.; DENG, M. A machine learning ensemble approach based on random forest and radial basis function neural network for risk evaluation of regional flood disaster: a case study of the Yangtze River Delta, China. **International Journal of Environmental Research and Public Health**, Basel, v. 17, n. 1, p. 49, 2020. DOI: <https://doi.org/10.3390/ijerph17010049>

DARMANIN, G. et al. Satellite-derived bathymetry for selected shallow Maltese coastal zones. **Applied Sciences**, Basel, v. 13, n. 9, e5238, 2023. DOI: <https://doi.org/10.3390/app13095238>

DUARTE, A. F. et al. A stochastic model of an early warning system for detecting anomalous incidence values of COVID-19. **Mathematical Geosciences**, Berlin, v. 56, n. 1, p. 41-54, 2024. DOI: <https://doi.org/10.1007/s11004-023-10096-4>.

ENQVIST, L. Rule-based versus AI-driven benefits allocation: GDPR and AIA legal implications and challenges for automation in public social security administration. **Information & Communications Technology Law**, Abingdon, v. 33, n. 2, p. 222-246, 2024. DOI: <https://doi.org/10.1080/13600834.2024.2349835>.

FROST, N. The impoverished publicness of algorithmic decision making. **Oxford Journal of Legal Studies**, Oxford, v. 44, n. 4, p. 780-807, 2024. DOI: <https://doi.org/10.1093/ojls/ggae027>.

GATZIOLIS, K. G.; TSELIKAS, N. D.; MOSCHOLIOS, I. D. Adaptive user profiling in e-commerce and administration of public services. **Future Internet**, Basel, v. 14, n. 5, e144, 2022. DOI: <https://doi.org/10.3390/fi14050144>.

GENG, R. *et al.* Analysis of the spatio-temporal dynamics of COVID-19 in Massachusetts via spectral graph wavelet theory. **IEEE Transactions on Signal and Information Processing over Networks**, Piscataway, v. 8, p. 670-683, 2022. DOI: <https://doi.org/10.1109/TSIPN.2022.3193252>.

GERRITS, L. Soul of a new machine: self-learning algorithms in public administration. **Information Polity**, Amsterdam, v. 26, n. 3, p. 237-250, 2021. DOI: <https://doi.org/10.3233/IP-200224>.

HAGEN, A.; ANDERSEN, T. M. Asset management, condition monitoring and digital twins: damage detection and virtual inspection on a reinforced concrete bridge. **Structure and Infrastructure Engineering**, Abingdon, v. 20, n. 7-8, p. 1242-1273, 2024. DOI: <https://doi.org/10.1080/15732479.2024.2311911>.

ISLAM, M. Z.; JILANI, M. M. A. K.; KARIM, M. R. Enhancing post-training evaluation of annual performance agreement training: a fusion of fsQCA and artificial neural network approach. **PLoS ONE**, San Francisco, v. 19, n. 6, e0305916, 2024. DOI: <https://doi.org/10.1371/journal.pone.0305916>.

JATO-ESPINO, D.; MAYOR-VITORIA, F. A statistical and machine learning methodology to model

rural depopulation risk and explore its attenuation through agricultural land use management. **Applied Geography**, Oxford, v. 152, p. 102870, 2023. DOI: <https://doi.org/10.1016/j.apgeog.2023.102870>.

KOWALSKI, R.; ESTEVE, M.; MIKHAYLOV, S. J. Improving public services by mining citizen feedback: an application of natural language processing. **Public Administration**, Oxford, v. 98, n. 4, p. 1011-1026, 2020. DOI: <https://doi.org/10.1111/padm.12656>.

MORELL-MONZÓ, S.; SEBASTIÁ-FRASQUET, M.-T.; ESTORNELL, J. Land use classification of VHR images for mapping small-sized abandoned citrus plots by using spectral and textural information. **Remote Sensing**, Basel, v. 13, n. 4, p. 681, 2021. DOI: <https://doi.org/10.3390/rs13040681>.

MOZGOVOY, V. Longitudinal estimation of stress-related states through bio-sensor data. **Applied Computing and Informatics**, Bingley, v. 21, n. 3-4, p. 186-197, 2025. DOI: <https://doi.org/10.1108/ACI-03-2021-0070>.

PENG, H. et al. Environmental risk and policy choices in an energy intensive region of China: an empirical study in Shanxi Province. **IEEE Access**, v. 8, p. 63134-63143, 2020.

ORGEIRA-CRESPO, P. *et al.* Decision algorithm for the automatic determination of the use of non-inclusive terms in academic texts. **Publications**, Basel, v. 8, n. 3, p. 41, 2020. DOI: <https://doi.org/10.3390/publications8030041>.

PENG, H.; LIU, S.; XING, Y.; YUE, X. Environmental risk and policy choices in an energy intensive region of China: an empirical study in Shanxi Province. **IEEE Access**, Piscataway, v. 8, p. 63134-63143, 2020. DOI: <https://doi.org/10.1109/ACCESS.2020.2984013>.

QU, S.; WANG, H.; HU, Z.; WANG, Z.; HU, S. Mixed-use urban land parcels identification integrating geospatial data and machine learning. **Geo-spatial Information Science**, Abingdon, v. 28, n. 4, p. 1743-1756, 2025. DOI: <https://doi.org/10.1080/10095020.2024.2374996>

REZAPOUR, M.; ELM SHAUSER, S. K. Artificial intelligence-based analytics for impacts of COVID-19 and online learning on college students' mental health. **PLOS ONE**, San Francisco, v. 17, n. 11, e0276767, 2022. DOI: <https://doi.org/10.1371/journal.pone.0276767>.

SALAZAR, A.; PÉREZ, J. F.; GALLEGU, J. VigIA: prioritizing public procurement oversight with machine learning models and risk indices. **Data & Policy**, Cambridge, v. 6, e75, 2024. DOI: <https://doi.org/10.1017/dap.2024.83>.

SIBITENDA, H.; GARBA, M.; ASANTE, B. S.; SSEGAWA, P. S. Extracting semantic topics about development in Africa from social media. **IEEE Access**, Piscataway, v. 12, p. 142343-142359, 2024. DOI: <https://doi.org/10.1109/ACCESS.2024.3466834>.

SOYLU, A. *et al.* Data quality barriers for transparency in public procurement. **Information**, Basel, v. 13, n. 2, 99, 2022. DOI: <https://doi.org/10.3390/info13020099>.

REZAPOUR, M.; ELM SHAUSER, S. K. Artificial intelligence-based analytics for impacts of COVID-19 and online learning on college students' mental health. **PLOS ONE**, San Francisco, v. 17, n. 11, e0276767, 2022. DOI: <https://doi.org/10.1371/journal.pone.0276767>.

SALAZAR, A.; PÉREZ, J. F.; GALLEGU, J. VigIA: prioritizing public procurement oversight with machine learning models and risk indices. **Data & Policy**, Cambridge, v. 6, e75, 2024. DOI: <https://doi.org/10.1017/dap.2024.83>.

SIBITENDA, H. *et al.* Extracting semantic topics about development in Africa from social media.

IEEE Access, Piscataway, v. 12, p. 142343-142359, 2024. DOI: <https://doi.org/10.1109/ACCESS.2024.3466834>.

QU, S. *et al.* Mixed-use urban land parcels identification integrating geospatial data and machine learning. **Geo-Spatial Information Science**, Wuhan, v. 28, n. 4, p. 1743-1756, 2025. DOI: <https://doi.org/10.1080/10095020.2024.2374996>

SOYLU, A. *et al.* Data quality barriers for transparency in public procurement. **Information**, Basel, v. 13, n. 2, art. 99, 2022. DOI: <https://doi.org/10.3390/info13020099>.

THARMAKULASINGAM, M. *et al.* TransAMR: an interpretable transformer model for accurate prediction of antimicrobial resistance using antibiotic administration data. **IEEE Access**, Piscataway, v. 11, p. 75337-75350, 2023. DOI: <https://doi.org/10.1109/ACCESS.2023.3296221>.

VALLE, L. M. *et al.* RegBR: a novel Brazilian government framework to classify and analyze industry-specific regulations. **PLOS ONE**, San Francisco, v. 17, n. 9, e0275282, 2022. DOI: <https://doi.org/10.1371/journal.pone.0275282>.

VENTURA, R. *et al.* Comparing the vibrational behavior of e-kick scooters and e-bikes: evidence from Italy. **International Journal of Transportation Science and Technology**, Shanghai, v. 16, p. 68-86, 2024. DOI: <https://doi.org/10.1016/j.ijtst.2023.10.010>.

VIDAÑA-VILA, E.; NAVARRO, J.; STOWELL, D.; ALSINA-PAGÈS, R. M. Multilabel acoustic event classification using real-world urban data and physical redundancy of sensors. **Sensors**, Basel, v. 21, n. 22, 7470, 2021. DOI: <https://doi.org/10.3390/s21227470>.

VOGL, T. M. *et al.* Smart technology and the emergence of algorithmic bureaucracy: artificial intelligence in UK local authorities. **Public Administration Review**, Washington, v. 80, n. 6, p. 946-961, 2020. DOI: <https://doi.org/10.1111/puar.13286>.

SOYLU, A. *et al.* Data quality barriers for transparency in public procurement. **Information**, Basel, v. 13, n. 2, 99, 2022. DOI: <https://doi.org/10.3390/info13020099>.

THARMAKULASINGAM, M. *et al.* TransAMR: an interpretable transformer model for accurate prediction of antimicrobial resistance using antibiotic administration data. **IEEE Access**, Piscataway, v. 11, p. 75337-75350, 2023. DOI: <https://doi.org/10.1109/ACCESS.2023.3296221>.

VALLE, L. M. *et al.* RegBR: a novel Brazilian government framework to classify and analyze industry-specific regulations. **PLOS ONE**, San Francisco, v. 17, n. 9, e0275282, 2022. DOI: <https://doi.org/10.1371/journal.pone.0275282>.

VENTURA, R. *et al.* Comparing the vibrational behavior of e-kick scooters and e-bikes: evidence from Italy. **International Journal of Transportation Science and Technology**, Amsterdã, v. 16, p. 68-86, 2024. DOI: <https://doi.org/10.1016/j.ijtst.2023.10.010>.

VIDAÑA-VILA, E. *et al.* Multilabel acoustic event classification using real-world urban data and physical redundancy of sensors. **Sensors**, Basileia, v. 21, n. 22, 7470, 2021. DOI: <https://doi.org/10.3390/s21227470>.

VOGL, T. M. *et al.* Smart technology and the emergence of algorithmic bureaucracy: artificial intelligence in UK local authorities. **Public Administration Review**, Washington, v. 80, n. 6, p. 946-961, 2020. DOI: <https://doi.org/10.1111/puar.13286>.

XIONG, W.; ZHOU, P. A search strategy for publications in interdisciplinary research. **Profesional de la información**, Barcelona, v. 32, n. 5, e320522, 2023. DOI: <https://doi.org/10.3145/epi.2023.sep.22>.

YEUNG, K. The new public analytics as an emerging paradigm in public sector administration. **Tilburg Law Review**, Tilburg, v. 27, n. 2, p. 1-32, 2022. DOI: <https://doi.org/10.5334/tilr.303>.

ZAIDANI, H. *et al.* Augmentation and classification of requests in Moroccan dialect to improve quality of public service: a comparative study of algorithms. **Future Internet**, Basileia, v. 17, n. 4, 176, 2025. DOI: <https://doi.org/10.3390/fi17040176>.

ZEKIĆ-SUŠAC, M.; MITROVIĆ, S.; HAS, A. Machine learning based system for managing energy efficiency of public sector as an approach towards smart cities. **International Journal of Information Management**, Amsterdã, v. 58, 102074, 2021. DOI: <https://doi.org/10.1016/j.ijinfomgt.2020.102074>.

ZHANG, Xiaowei; ZHANG, Yueyan. Research on the problems and reform strategies of government statistics based on big data technology. **Applied Mathematics and Nonlinear Sciences**, Varsóvia, v. 9, n. 1, 2024. ISSN 2444-8656.

ZHAO, X.; WANG, S.; WANG, H. Organizational geosocial network: a graph machine learning approach integrating geographic and public policy information for studying the development of social organizations in China. **ISPRS International Journal of Geo-Information**, Basileia, v. 11, n. 5, 318, 2022. DOI: <https://doi.org/10.3390/ijgi11050318>.