

APLICAÇÃO DE *MACHINE LEARNING* NA GESTÃO DE RISCO DE QUEDA DE ÁRVORES URBANAS EM MEDIANEIRA, PARANÁ

APPLICATION OF MACHINE LEARNING IN MANAGING URBAN TREE FALL RISK IN MEDIANEIRA, PARANÁ

Allan Rodrigo Nunho dos Reis¹, Severo Ivasko Junior², Daniel Zambiazzi Miller³

RESUMO

A avaliação de risco arbóreo urbano é uma atividade essencial, porém complexa e morosa, o que dificulta a gestão eficiente das florestas urbanas. Esta pesquisa teve como objetivos desenvolver um modelo preditivo de *Machine Learning* (ML) para a classificação de risco de árvores e identificar as variáveis mais relevantes associadas à probabilidade de queda. Foram utilizados dados de 1.141 árvores de rua do município de Medianeira, Paraná, contemplando 33 variáveis dendrométricas, fitossanitárias e do entorno. O modelo foi implementado por meio do algoritmo *XGBoost*, otimizado pelo *GridSearchCV*, com abordagem híbrida de *oversampling* (SMOTE) e ponderação de classes. O modelo apresentou acurácia de 71% e alta precisão (94%) para a classe “Ruim”, evidenciando seu potencial como ferramenta de triagem. A análise de importância das variáveis indicou que combinações de sintomas fitossanitários — como galhos secos, presença de fungos e desequilíbrio de copa — são mais preditoras do risco do que variáveis dendrométricas isoladas. Conclui-se que o uso de ML na arborização urbana aprimora o diagnóstico e subsidia decisões técnicas baseadas em dados objetivos.

Palavras-chave: Avaliação de Risco; Floresta Urbana; Inteligência Artificial; *XGBoost*.

ABSTRACT

Urban tree risk assessment is an essential yet complex and time-consuming activity, hindering the efficient management of urban forests. This search aimed to develop a predictive Machine Learning (ML) model for classifying tree risk and identifying the most relevant variables associated with the probability of tree failure. Data from 1,141 street trees in the municipality of Medianeira, Paraná, were used, encompassing 33 dendrometric, phytosanitary, and environmental variables. The model was implemented using the *XGBoost* algorithm, optimized through *GridSearchCV*, with a hybrid oversampling approach (SMOTE) and class weighting. The model demonstrated 71% accuracy and high precision (94%) for the "Poor" class, demonstrating its potential as a screening tool. Variable importance analysis indicated that combinations of phytosanitary symptoms—such as dry branches, presence of fungi, and canopy imbalance—are more predictive of risk than isolated dendrometric variables. It is concluded that the use of ML in urban forest improves diagnosis and supports technical decisions based on objective data.

Keywords: Urban Forestry; Artificial Intelligence; *XGBoost*; Risk Assessment; Machine Learning.

Recebido em 29.10.2026 e aceito em 06.05.2026

1 Engenheiro Florestal. Doutor. Prefeitura Municipal de Pinhais/PR. Curitiba/PR. Email: allan.nunho@gmail.com

2 Engenheiro Florestal. Doutor. Prefeitura Municipal de Apiúna/SC. Apiúna/SC. Email: severoivasko@gmail.com

3 Engenheiro Florestal. Mestre. Universidade Federal do Paraná. Morretes/PR. Email: daniel.z.miller@hotmail.com

INTRODUÇÃO

As árvores urbanas oferecem benefícios ambientais e sociais essenciais relacionados à melhoria microclimática, fixação de carbono e bem-estar psicossocial, além de outros. Contudo, para que estas árvores possam prover tais benefícios, elas precisam estar em boas condições fitossanitárias. Como medida para esta questão, nos dias atuais existe uma preocupação crescente quanto ao risco de ruptura das árvores situadas em áreas urbanas, uma vez que em caso de queda, elas podem ocasionar danos a pessoas e propriedades (JODAS et al., 2024; ISA et al., 2025).

Por isso, diversas pesquisas têm buscado estruturar indicadores que sejam associados a defeitos e consequentes rupturas das árvores, tais como a deterioração e uniões fracas dos galhos, cavidades, rachaduras, danos nas raízes, e presença de pragas e doenças (JAHANI; SAFFARIHA, 2022; MARIA et al., 2023; JODAS et al., 2024; SRIVANIT; KAEWKHOW, 2024; ISA et al., 2025). A junção destes fatores compõe a avaliação completa dos riscos apresentados por uma árvore, a qual inclui inspeções visuais à procura de sinais biomecânicos e estruturais que comprometam a segurança da árvore e do seu entorno (CAGGIU et al., 2023; SRIVANIT; KAEWKHOW, 2024).

Porém, o manejo das árvores urbanas é desafiador, especialmente quanto à avaliação de risco de queda, pois os métodos convencionais baseados em inspeções visuais são subjetivos, lentos e custosos (CAGGIU et al., 2023; MARIA et al., 2023). Além disso, os dados de campo existentes ainda não são capazes de caracterizar com exatidão atributos vitais sobre a saúde das árvores, incluindo defeitos e riscos potenciais, o que é fundamental para um monitoramento eficaz que otimize o uso de recursos (ISA et al., 2025). Tal cenário é agravado pela uma dificuldade de interpretação e detecção de padrões complexos pelos métodos tradicionais, devido à subjetividade.

Embora os métodos convencionais de inspeção visual carreguem um componente de subjetividade, eles representam a sistematização do conhecimento especialista acumulado. O emprego de Machine Learning sobre esses dados não visa apenas replicar as tendências do avaliador, mas sim processar essas percepções para identificar padrões biofísicos consistentes. Dessa forma, o modelo atua transformando observações qualitativas em indicadores preditivos robustos, mitigando inconsistências individuais e otimizando a detecção de riscos reais (VALENCIA-ARIAS et al., 2025). Assim, a avaliação do risco associado à ruptura de árvores urbanas representa um problema proeminente nas cidades, o que demanda um aprimoramento do cuidado e do monitoramento das árvores urbanas, a fim de superar as incertezas e auxiliar nas tomadas de decisão (CAGGIU et al., 2023; VALENCIA-ARIAS et al., 2025).

Porém, o manejo das árvores urbanas é desafiador, especialmente quanto à avaliação de risco de queda, pois os métodos convencionais baseados em inspeções visuais são subjetivos, lentos e custosos (CAGGIU et al., 2023; MARIA et al., 2023). Contudo, embora tais métodos carreguem um componente de subjetividade, eles representam a sistematização do conhecimento especialista acumulado. O emprego de *Machine Learning* sobre esses dados

não visa apenas replicar as tendências do avaliador, mas sim processar essas percepções para identificar padrões biofísicos consistentes.

Dessa forma, o modelo atua transformando observações qualitativas em indicadores preditivos robustos, mitigando inconsistências individuais e otimizando a detecção de riscos reais (VALENCIA-ARIAS et al., 2025). Essa abordagem permite superar as incertezas inerentes aos métodos tradicionais, auxiliando nas tomadas de decisão e aprimorando o monitoramento contínuo das cidades (CAGGIU et al., 2023; ISA et al., 2025).

Neste sentido, tem-se testado ferramentas que auxiliam na gestão arbórea, tais como os modelos preditivos, os quais podem indicar com maior precisão os fatores que se relacionam com o risco de queda destas árvores (MARIA et al., 2023; SRIVANIT; KAEWKHOW, 2024). Nos últimos anos, aumentaram os investimentos no desenvolvimento de modelos que utilizam técnicas de Inteligência Artificial (IA), notadamente de *Machine Learning* (ML), ou aprendizado de máquina, para prever o risco de queda de árvores urbanas. As técnicas de ML se destacam por permitir a análise de elevados volumes de dados, como um complemento às medidas tradicionais, o que possibilita a identificação de padrões e gera avaliações mais precisas e dinâmicas quanto ao risco de queda (RYAN; MASSARON, 2025; VALENCIA-ARIAS et al., 2025). Além disso, as técnicas de ML favorecem a otimização de recursos, a priorização de inspeções, o aumento da segurança, e o auxílio à gestão pública nas ações de manejo arbóreo.

Contudo, o conhecimento sobre como as aplicações de ML na gestão de risco de árvores urbanas contribuem para a avaliação de risco de queda ainda são incipientes, sobretudo no Brasil. Com este objetivo, Srivanit e Kaewkhow (2024) integraram um Sistema de Informações Geográficas (SIG) com um modelo de ML gerado a partir do algoritmo CHAID para avaliar 3659 árvores na Tailândia, e alcançaram 87,35% de precisão na classificação de risco de queda, explicada principalmente pela presença de cavidades e pragas. Jahani e Saffariha (2022) compararam modelos de redes neurais e de *Support Vector Machine* (SVM), uma técnica de ML, para prever o risco de queda de árvores da espécie *Platanus orientalis* L. em cidades do Iran e concluíram que o SVM foi o mais preciso (97,5%).

No Brasil, Maria et al. (2023) conseguiram otimizar a avaliação de risco de queda em 230 árvores urbanas em Itanhaém, São Paulo, usando o algoritmo C4.5, diminuindo de 36 variáveis visuais para apenas 14, com 73% de acerto. Já Jodas et al. (2024) propuseram o uso de técnicas de ML para a predição de risco de queda de árvores da cidade de São Paulo, e constataram que o modelo de *stacking generalization* alcançou o menor erro médio absoluto.

Porém, a aplicação do algoritmo de *ensemble XGBoost* (*eXtreme Gradient Boosting*) ainda não foi relatada na literatura especializada para a predição do risco de queda de árvores urbanas. Assim, esta pesquisa teve como objetivo testar o algoritmo *XGBoost* para desenvolver um modelo preditivo baseado em *Machine Learning*, visando otimizar a avaliação de risco de árvores urbanas em Medianeira, Paraná.

MATERIAL E MÉTODOS

Área de estudo

A pesquisa foi conduzida na área urbana do município de Medianeira, localizado na macrorregião oeste do estado do Paraná, nas proximidades dos municípios de Cascavel e Foz do Iguaçu (Figura 1). O território de Medianeira se estende por uma área de 328,732 km², sendo que 17,597 km² compreendem o perímetro urbano, o qual é dividido em duas regiões: Sede (31,72 km²) e Distrito Maralúcia (0,78 km²).

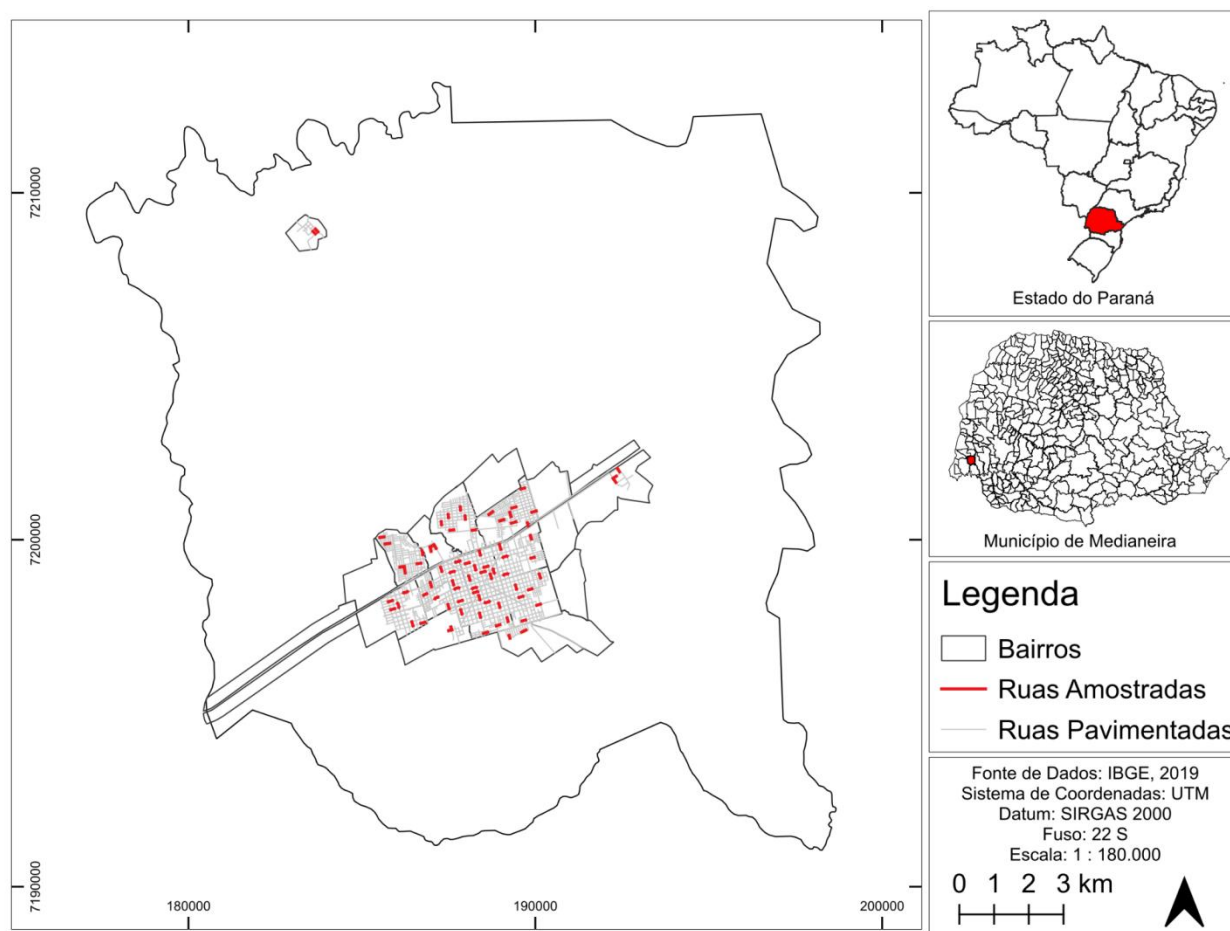


Figura 1. Localização do município de Medianeira, PR.
Figure 1. Location of the municipality of Medianeira, PR.

Em consulta aos relatórios de ocorrência de eventos extremos registrados em Medianeira, por meio do sistema GeoDC da Coordenadoria Estadual de Proteção e Defesa Civil do Paraná (DEFESA CIVIL DO PARANÁ, 2023), e consulta ao 9º Grupamento de Bombeiros – Foz do Iguaçu, quanto os seguintes eventos: Tempestade Local/Convectiva, Ondas de Frio, Ondas de Calor, Incêndio Florestal, Ciclones, Poda ou Corte de árvores, para o município de Medianeira. A partir destas consultas, foi verificado que durante o período de 2017 até 2023 foram registrados 40 eventos de corte ou poda de árvores, quatro vendavais, dois eventos de chuva intensa e um de granizo.

Procedimentos metodológicos

O *dataset* utilizado foi referente às coletas em campo para o diagnóstico arbóreo da arborização de ruas de Medianeira para a elaboração do Plano Municipal de Arborização Urbana (PMAU) do município. Os dados consistiram no levantamento de variáveis relativas a informações ecológicas, dendrométricas e fitossanitárias, e do entorno de cada indivíduo arbóreo amostrado. A seleção das parcelas amostrais foi feita por trechos de rua, conforme metodologia proposta por Maria, Biondi e Zamproni (2017). Assim, ao todo, foram estabelecidas 77 parcelas amostrais.

Os dados foram coletados com auxílio do aplicativo para *smartphones* de código aberto ODK Collect, o qual dispõe de armazenamento dos dados em nuvem e a possibilidade de coletar as coordenadas geográficas. As informações ecológicas consistiram na identificação de cada indivíduo amostrado. A identificação botânica foi realizada *in loco* pela equipe técnica. Nos casos em que a identificação imediata não foi possível, utilizaram-se os aplicativos *PlantNet*[®] e *LeafSnap*[®] como ferramentas auxiliares ou, de forma complementar, coletaram-se amostras para posterior análise por especialistas. Para a mensuração da Circunferência à Altura do Peito (CAP) e do Diâmetro de Copa (DC) foram utilizadas trenas métricas.

A altura total foi estimada visualmente, enquanto as alturas de bifurcação e de início de copa foram mensuradas com trenas métricas. As condições estruturais e fitossanitárias dos indivíduos amostrados foram avaliadas visualmente, observando-se o sistema radicular visível, o tronco e a copa, à procura de fatores que pudessem representar expressivamente condições de risco à saúde e estabilidade da planta.

As espécies foram classificadas como nativa, com ocorrência natural no território brasileiro, exótica, introduzida no território brasileiro, e exótica invasora, com potencial de invadir áreas naturais no país. Os nomes científicos aceitos foram verificados na plataforma Flora do Brasil (REFLORA, 2025), considerando o sistema APG IV (2016). As espécies exóticas invasoras foram consultadas na Lista de Espécies Exóticas Invasoras do Paraná, instituída pela Portaria IAP nº 59, de 15 de abril de 2015 (INSTITUTO AMBIENTAL DO PARANÁ (IAP), 2015), e na base de dados do Instituto Hórus (INSTITUTO HÓRUS, 2023).

Os 1.141 indivíduos de porte arbóreo amostrados se distribuíram em 110 espécies e 42 famílias botânicas. A família com maior riqueza de espécies na arborização das ruas de Medianeira é Fabaceae, com 13 espécies, o que corresponde a 12,03% do total; seguida por Arecaceae, com 10 espécies (9,26%); Bignoniaceae, com 9 espécies (8,33%) e Myrtaceae, com 7 espécies (6,48%).

Quanto ao hábito das espécies amostradas, 63,89% são classificadas como árvores, 25% como arbustos, 9,26% como palmeiras, e 1,85% de outros hábitos de vida (como as yucas e os mamoeiros). Com relação à origem, das 108 espécies que puderam ser identificadas, 47 são nativas do Brasil (43,52%) e 61 são exóticas (56,48%). Dentre as

espécies exóticas, 16 espécies (14,81%) são consideradas invasoras no estado do Paraná (PARANÁ, 2015), tendo o seu plantio proibido na arborização urbana.

Na avaliação do meio físico, foram coletadas as informações de localização como nome da rua e bairro, lado par ou ímpar da calçada, número da residência mais próxima e coordenadas do indivíduo. Além disso, foram mensuradas a posição da árvore na calçada, distância ao recuo do lote, distância da guia/meio fio, área do canteiro, distância e tipo da fixação aérea, e distância até a árvore seguinte.

Em campo, foi utilizada a classificação fitossanitária proposta por Milano (1984), a qual é composta pelas seguintes categorias:

a) Boa – com ausência de sinais de pragas, doenças ou injúrias mecânicas, apresenta a forma característica da espécie e não demanda medidas de correção;

b) Satisfatória – com condição física e vigor medianos, pode ter sofrido podas pesadas, mas conseguiu se reestabelecer satisfatoriamente, ou ainda demanda reparos de danos físicos ou controle de pragas ou doenças;

c) Ruim – com estado geral de declínio de vitalidade, muitos danos físicos, ataque de pragas ou doenças, tortuosidade, poda intensa que descaracteriza a espécie e prejudica sua recuperação, demandando muitas medidas de correção;

d) Árvore morta – com aparente ausência de atividades fisiológicas.

Contudo, para as análises desta pesquisa, optou-se por unificar as classes “Ruim” e “Morta” na classe “Ruim”, devido à pouca quantidade de árvores mortas encontradas. Esta classificação foi utilizada para alimentar o modelo preditivo, de modo a, em um primeiro momento, verificar se o algoritmo de ML alcançaria uma acurácia satisfatória na classificação das árvores conforme havia sido feito anteriormente pelos técnicos.

A Figura 2 apresenta exemplos de árvores incluídas nas quatro classes de condições fitossanitárias.



Figura 2. Exemplos de árvores encontradas na arborização de ruas de Medianeira, PR. Classes fitossanitárias: A: Boa; B: Satisfatória; C: Ruim; D: Morta.

Figure 2. Examples of trees found in the streetscapes of Medianeira, PR. Phytosanitary classes: A: Good; B: Satisfactory; C: Poor; D: Dead.

Realizou-se um processo de *data wrangling* inicial para padronização e exclusão de inconsistências. A variável-alvo “condição geral” foi agrupada em três categorias: “Boa”, “Satisfatória” e “Ruim”. Os dados foram divididos em 70% para treinamento e 30% para teste, de forma estratificada.

O modelo foi construído com o algoritmo *XGBoost* (*eXtreme Gradient Boosting*), com otimização de hiperparâmetros pelo *GridSearchCV* (RYAN; MASSARON, 2025). Aplicou-se a técnica de *oversampling* SMOTE para balanceamento e ponderação de classes durante o treinamento. As métricas de avaliação foram *Accuracy*, *Precision*, *Recall* e *F1-Score*, recomendadas por Van der Plas (2023) e por Ryan e Massaron (2025).

Após o treinamento do modelo final, a importância de cada variável foi extraída por meio da propriedade *feature_importances*, embutida no algoritmo *XGBoost* (RYAN; MASSARON, 2025). As variáveis foram ranqueadas conforme sua contribuição para a redução do erro preditivo, visando interpretar os fatores mais determinantes de risco.

Todas as análises foram conduzidas em ambiente Python 3.11, utilizando as bibliotecas *Pandas*, *Scikit-learn*, *Imbalanced-learn*, *XGBoost*, *Matplotlib* e *Seaborn*, por meio da plataforma *Google Colaboratory* (Colab).

RESULTADOS E DISCUSSÃO

A acurácia geral de 71% e o *F1-Score* de 72% são reflexos da natureza qualitativa da base de dados. A metodologia de Milano (1984) fundamenta-se na percepção visual de defeitos, o que introduz uma carga de subjetividade na rotulagem das classes, especialmente nas categorias intermediárias. Todavia, essa moderação na acurácia global é compensada pela especificidade do modelo: ao atingir 94% de precisão na classe “Ruim”, o algoritmo prova ser capaz de transcender a subjetividade humana nos casos críticos, validando sua aplicação como uma ferramenta de triagem confiável para a gestão de riscos.

Tabela 1. Indicadores de performance do modelo preditivo.

Table 1. Predictive model performance indicators.

Classe	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	Support (N° de árvores no teste)
Boa	0,82	0,78	0,80	209
Ruim	0,94	0,48	0,64	33
Satisfatória	0,51	0,64	0,57	101
Acurácia geral			0,71	343
Média macro	0,76	0,64	0,67	343
Média ponderada	0,74	0,71	0,72	343

O principal destaque do modelo foi a sua elevada precisão (*precision*) de 94% para a classe “Ruim”. Isso indica que, quando o modelo classificou uma árvore como sendo de alto risco, ele estava correto em 94% dos casos. Na prática, isso transforma o modelo em uma

ferramenta de triagem com elevada confiabilidade, pois as árvores que ele apontou como problemáticas quase certamente necessitam de uma inspeção urgente, minimizando o desperdício de tempo da equipe de campo com falsos positivos.

Embora o conjunto de dados inclua palmeiras e arbustos, que possuem estruturas biomecânicas distintas das árvores, a alta precisão do modelo (94% para a classe 'Ruim') indica que os sinais fitossanitários críticos (fungos, cavidades e necrose) são indicadores universais de declínio que o algoritmo foi capaz de capturar, independentemente do hábito de crescimento. Nestas situações, o técnico responsável pela avaliação das árvores deve aconselhar a adoção de medidas de mitigação e o agendamento do tratamento o mais breve possível, sobretudo se houver um alvo de alto valor e a falha for iminente (MARIA et al., 2023; SRIVANIT; KAEWKHOW, 2024).

Embora a sensibilidade (*recall*) para a classe “Ruim” tenha sido de 48%, indicando que o modelo não identificou a totalidade dos indivíduos em risco severo, esse comportamento é compensado pela elevada precisão (94%) obtida. Na gestão pública, esse equilíbrio é estratégico: o modelo atua como um sistema de alerta proativo e de elevada confiabilidade, garantindo que as intervenções priorizadas sejam aplicadas em alvos que efetivamente demandam manejo, otimizando recursos e evitando falsos positivos que poderiam levar a remoções desnecessárias. Portanto, o modelo não substitui a inspeção completa, mas serve como uma ferramenta de triagem proativa para identificar metade das árvores críticas com erro quase nulo.

Para a classe "Boa", o modelo apresentou um desempenho satisfatório, com *F1-Score* de 80%, identificando corretamente a maioria das árvores saudáveis. A dificuldade do modelo em classificar a categoria “Satisfatória” reside na sobreposição de atributos (*feature overlap*) com as classes adjacentes. Por ser uma condição intermediária, seus sinais fitossanitários são menos distintos, o que gera incerteza algorítmica. Embora a técnica de *oversampling SMOTE* tenha sido aplicada para equilibrar as proporções das classes, ela atua sobre a distribuição numérica, mas não elimina a ambiguidade qualitativa intrínseca ao diagnóstico visual. Assim, o modelo tende a ser mais assertivo nos extremos (Boa ou Ruim), onde os padrões biológicos são mais evidentes, e mais conservador na zona de transição. Nestas situações o manejo recomendado, como podas leves ou endoterapia possui menor urgência de segurança pública (SRIVANIT; KAEWKHOW, 2024).

Para estas árvores, pode ser importante realizar medidas de manutenção ou mitigação para essas árvores, mas não é uma necessidade urgente. Conforme Srivanit e Kaewkhow (2024), nestes casos são recomendadas medidas de mitigação como podas e aplicação de técnicas de endoterapia.

Apesar de a acurácia geral ter sido de 71%, conforme destacado por Maria et al. (2023), o valor do modelo não está em classificar perfeitamente todas as árvores, mas sim em reduzir drasticamente o espaço de busca para a inspeção humana. Além disso, ao direcionar a atenção dos técnicos para um grupo de árvores com 94% de chance de realmente

estarem em risco, o modelo otimiza recursos e aumenta a eficiência da gestão preventiva da arborização urbana. Assim, a partir dos resultados de classificação de risco arbóreo, pode-se desenvolver um sistema que além de alertar sobre árvores ruins, também otimiza as rotas de inspeção e os cronogramas de manejo arbóreo (JODAS et al., 2024; ISA et al., 2025).

Com relação às variáveis preditivas, a mais importante indicada pelo modelo de ML foi a combinação de problemas presente na copa, representados por galhos secos, fungos e desequilíbrio (Figura 3). Assim, verifica-se que não apenas um galho seco indica risco, mas sim a co-ocorrência destes três sintomas. O mesmo vale para a presença de fungo nas raízes e a combinação de união fraca, lesão e cavidade no tronco. Esta combinação de variáveis já foi relatada em outras pesquisas (MARIA et al., 2023; JODAS et al., 2024; SRIVANIT; KAEWKHOW, 2024).

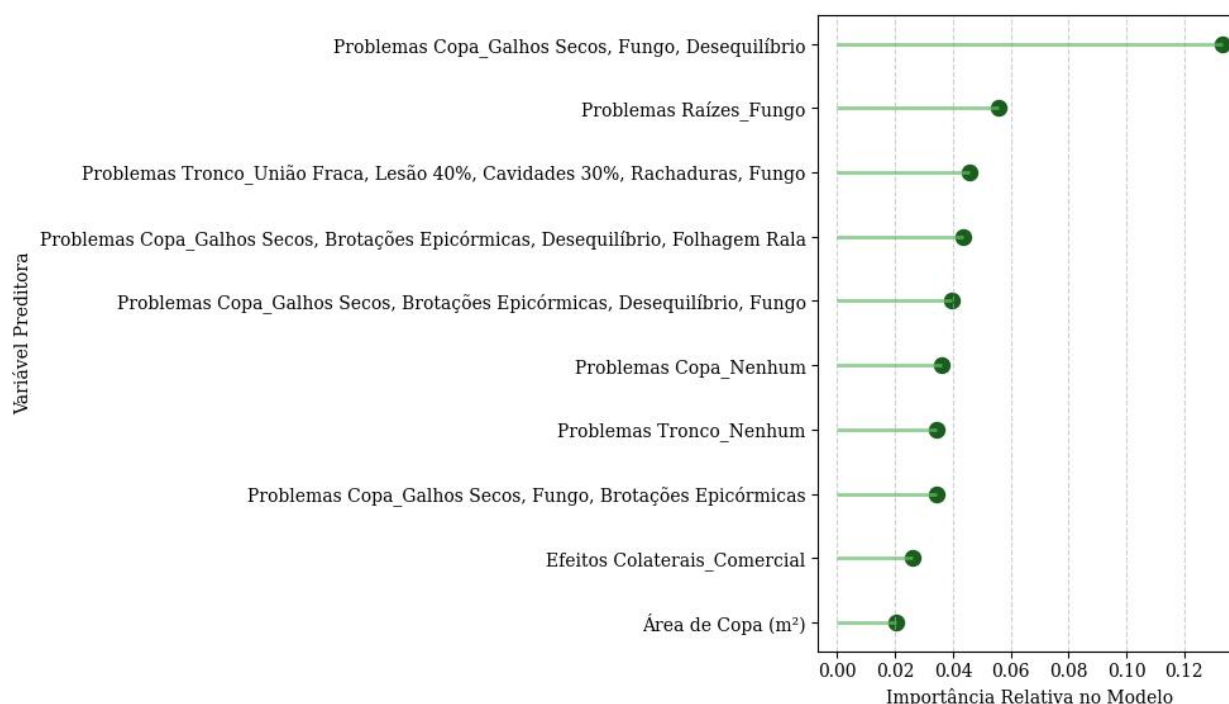


Figura 3. *Ranking* das principais variáveis preditoras do risco de queda das árvores de rua de Medianeira, Paraná.

Figure 3. Ranking of the main predictive variables of the risk of street trees falling in Medianeira, Paraná.

As quatro variáveis mais importantes foram descrições de problemas fitossanitários. A primeira variável puramente de medição, a área de copa (m²), só apareceu na 10^o colocação. Isso reforça a constatação de que os defeitos visíveis possuem maior capacidade preditiva de risco do que as medidas de dimensão das árvores. Maria et al. (2023) verificaram a combinação de defeitos estruturais na copa como mais determinante para o risco do que variáveis individuais em árvores urbanas. Srivanit e Kaewkhaw (2024) também identificaram fatores ligados à copa, tais como cavidades, infestações de pragas, danos mecânicos, galhos mortos e crescimento epicórmico, como as mais importantes para o risco de queda. Assim, reforça-se que apenas o fato de a árvore ter grandes dimensões, ou estar inclinada, por si só, não caracteriza o risco relevante de queda.

O *ranking* de importância das variáveis gerado pelo modelo de ML otimiza as avaliações de campo, especialmente para técnicos municipais que possuem elevados volumes de árvores para avaliar, ao direcionar o foco para os defeitos mais críticos e tornar a inspeção mais rápida e direcionada. Além disso, os resultados reforçam a literatura sobre modelagem preditiva com ML em dados tabulares. Conforme Ryan e Massaron (2025), o algoritmo *XGBoost* é altamente eficaz na classificação de dados estruturados e permite explorar interações complexas entre variáveis. Somado a isso, a técnica de *feature importance* aplicada ao modelo permitiu identificar padrões não triviais, valorizando a sua interpretação. Este aspecto é destacado como vantagem dos modelos baseados em ML (VAN DER PLAS, 2023).

Ao revelar que os principais preditores envolvem a combinação de sintomas visuais, os resultados oferecem subsídios práticos para gestores públicos, sugerindo que protocolos tradicionais podem ser otimizados com foco em um conjunto reduzido de variáveis críticas. Essa abordagem contribui para a racionalização de tempo e recursos, ao mesmo tempo em que mantém uma elevada precisão na identificação de árvores com risco iminente, alinhando-se às recomendações de uso estratégico da IA na gestão de árvores urbanas (MARIA et al., 2023).

Além disso, o modelo desenvolvido serve como um complemento à Norma NBR 16246 - Florestas urbanas — Manejo de árvores, arbustos e outras plantas lenhosas - Parte 3: Avaliação de risco de árvores (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT), 2013). Neste sentido, ele funciona principalmente como um aprimoramento do "Nível 1", de avaliação expedita, mas também como uma base para a avaliação de "Nível 2", ao indicar as variáveis prioritárias a serem consideradas.

Assim, os resultados podem orientar protocolos de inspeção focados nas variáveis de maior importância, reduzindo a subjetividade, os custos e o tempo de avaliação. Além disso, a abordagem pode ser integrada a sistemas inteligentes de priorização de manejo e planejamento preventivo em escala municipal, tais como plataformas *web*.

Para futuras pesquisas, recomenda-se explorar o motivo das combinações de defeitos situados na copa serem tão importantes para o risco de queda de árvores urbanas, buscando-se investigar a existência de sinergias biomecânicas ou fitopatológicas que expliquem esta interação. Por fim, como o modelo foi treinado com dados coletados na cidade de Medianeira, a qual possui suas próprias características urbanas, aplicações em outras cidades podem necessitar de ajustes ou validação que considerem diferentes espécies, clima e/ou condições urbanas.

CONCLUSÕES

O modelo desenvolvido com o algoritmo *XGBoost* apresentou desempenho promissor, destacando-se pela elevada confiabilidade (*precision*) na identificação de árvores em risco. Embora a sensibilidade (*recall*) para essa classe tenha sido moderada, a capacidade do

modelo de gerar alertas com poucos falsos positivos o qualifica como uma ferramenta eficaz de triagem.

O *ranking* gerado por *Machine Learning* possibilitou identificar os fatores determinantes do risco e definir um conjunto otimizado de variáveis-chave para os protocolos de avaliação em campo, tornando-os mais ágeis, eficientes e adaptáveis a diferentes realidades urbanas. Essa abordagem fortalece a tomada de decisão técnica, pautada em dados objetivos.

Além de otimizar os protocolos de inspeção, o modelo apresenta potencial de integração com sistemas digitais e georreferenciados, favorecendo a identificação de áreas críticas e o planejamento estratégico de ações preventivas. Sua capacidade de replicação em diferentes contextos urbanos — desde que apoiado em inventários arbóreos consistentes — amplia o alcance da metodologia e contribui para a padronização de processos em escala regional ou nacional, fortalecendo políticas públicas de manejo sustentável e qualificando a tomada de decisão baseada em evidências.

AGRADECIMENTOS

À Prefeitura Municipal de Medianeira, Paraná, pela disponibilização da base de dados e à Universidade Livre do Meio Ambiente (UNILIVRE), pela concessão da bolsa de pesquisa.

REFERÊNCIAS

Angiosperm Phylogeny Group (APG). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. **Botanical Journal of the Linnean Society**, Oxford, v. 181, n. 1, p. 1-20, 2016.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). **NBR 16246-3**: Florestas urbanas – Manejo de árvores, arbustos e outras plantas lenhosas - Parte 3: Avaliação de risco de árvores. Rio de Janeiro: ABNT, 2013b. 20 p.

CAGGIU, L.; FIORANI, F.; CORRADINI, E.; FELICE, E.; MINELLI, A. Spatial Analysis of Risk Exposure of Urban Trees: A Case Study from Bologna (Italy). **Urban Science**, Basel, v. 7, n. 4, 123, 2023.

DEFESA CIVIL DO PARANÁ. **GeoDC - Coordenadoria Estadual de Proteção e Defesa Civil do Paraná** - Mapa Público. 2023. Disponível em: <https://geodc.geo.pr.gov.br/geodc/pages/templates/initial_public.jsf?windowId=ce0>. Acesso em: 20 set. 2023.

INSTITUTO AMBIENTAL DO PARANÁ (IAP). **Portaria IAP nº 059, de 15 de abril de 2015**. Reconhece a Lista Oficial de Espécies Exóticas Invasoras para o Estado do Paraná. Curitiba, 2015. Disponível em: <https://www.iat.pr.gov.br/sites/agua-terra/arquivos_restritos/files/documento/2020-04/folder_web_geral.pdf>. Acesso em 10 set. 2023.

INSTITUTO HÓRUS. 2023. **Base de Dados Nacional de Espécies Exóticas Invasoras**. 2023. Disponível em: <<https://bd.institutohorus.org.br/>>. Acesso em 20 set. 2023.

ISA, M. M.; ZAINAL, M. H.; ZAKARIA, M. A.; TAHAR, K. N.; ZHUANG, Q. Utilizing Tree Risk Assessment (TRA) and Unmanned Aerial Vehicle (UAV) as a pre-determine tree hazard

identification. **Environment-Behaviour Proceedings Journal**, Shah Alam, v. 10, n. 32, p. 359-366, 2025.

JAHANI, A.; SAFFARIHA, M. Tree failure prediction model (TFPM): machine learning techniques comparison in failure hazard assessment of *Platanus orientalis* in urban forestry. **Natural Hazards**, Berlin, v. 110, n. 2, p. 881-898, 2022.

JODAS, D. S.; BRAZOLIN, S.; VELASCO, G. D. N.; LIMA, R. A.; YOJO, T.; PAPA, J. P. Urban tree failure probability prediction based on dendrometric aspects and machine learning models. **Computers, Environment and Urban Systems**, Amsterdam, v. 108, n. 102074, 2024.

MARIA, T. R. B. C.; BIONDI, D.; BEHLING, A.; REIS, A. R. N.; ZAMPRONI, K.; HO, T. L. Application of artificial intelligence for tree risk assessment optimization in Itanhaém – São Paulo, Brazil. **Urban Forestry & Urban Greening**, Amsterdam, v. 81, p. 01-09, 2023.

MARIA, T. R. B. C.; BIONDI, D.; ZAMPRONI, K. Proposta de calçadas acessíveis em consonância com a arborização viária estabelecida em Itanhaém – SP. **Revista da Sociedade Brasileira de Arborização Urbana**, Piracicaba, v. 12, n. 4, p. 87-101, 2017.

MILANO, M S. **Avaliação e Análise da Arborização de Ruas de Curitiba-PR**. 130f. Dissertação (Mestrado em Ciências Florestais) – Universidade Federal do Paraná, Curitiba, Paraná, 1984.

REFLORA. **Flora e Funga do Brasil**. 2023. Jardim Botânico do Rio de Janeiro. Disponível em: <<http://floradobrasil.jbrj.gov.br/>>. Acesso em: 28 set. 2023.

RYAN, M.; MASSARON, L. **Machine Learning for Tabular Data** – XGBoost, Deep Learning, and AI. Shelter Island: Manning Publications Co., 2025. 504p.

SRIVANIT, M.; KAEWKHOW, S. A machine learning-based protocol to support visual tree assessment and risk of failure classification on a university campus. **Urban Forestry & Urban Greening**, Amsterdam, v. 99, p. 128420, 2024.

VALENCIA-ARIAS, A.; GARCIA, J. A. J.; AGUDELO-CEBALLOS, E.; LEÓN, A. J. A. O.; ROJAS, E. M.; HENRÍQUEZ, J. L.; RAMÍREZ-RAMÍREZ, D. M. Machine learning applications in risk management: Trends and research agenda. **F1000Research**, Londres, v. 14, n. 233, 2025.

VAN DER PLAS, J. **Python Data Science Handbook: Essential Tools for Working with Data**. Sebastopol: O'Reilly, 2023. 591p.