

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Exatas
Programa de Pós-Graduação em Ciência da Computação

Gabriel de Oliveira Campos Pacheco

Análise espacial e mineração de dados de arborização urbana em Belo
Horizonte

Belo Horizonte
2020

Gabriel de Oliveira Campos Pacheco

Análise espacial e mineração de dados de arborização urbana em Belo Horizonte

Versão Final

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Ciência da Computação.

Orientador: Clodoveu Augusto Davis Júnior

Belo Horizonte
2020

© 2020, Gabriel de Oliveira Campos Pacheco.
. Todos os direitos reservados

Ficha catalográfica elaborada pela bibliotecária Belkiz Inez Rezende Costa
CRB 6ª Região nº 1510

Pacheco, Gabriel de Oliveira Campos.

P116a Análise espacial e mineração de dados de arborização urbana em Belo Horizonte / Gabriel de Oliveira Campos Pacheco. — Belo Horizonte, 2020.
xxii, 64 f. il.; 29 cm.

Dissertação (mestrado) - Universidade Federal de Minas Gerais – Departamento de Ciência da Computação
Orientador: Clodoveu Augusto Davis Júnior.

1. Computação – Teses. 2. Arborização das cidades. 3. Mineração de dados (Computação). 4. Banco de dados geográficos. I.Orientador. II.Título.

CDU 519.6*73(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FOLHA DE APROVAÇÃO

Análise espacial e mineração de dados de arborização urbana em Belo Horizonte

GABRIEL DE OLIVEIRA CAMPOS PACHECO

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

PROF. CLODOVEU AUGUSTO DAVIS JÚNIOR - Orientador
Departamento de Ciência da Computação - UFMG

DRA. KARLA ALBUQUERQUE DE VASCONCELOS BORGES
Empresa de Informática e Informação do Município de Belo Horizonte - PRODABEL

PROF. JUGURTA LISBOA FILHO
Departamento de Informática - UFV

PROFA. MIRELLA MOURA MORO
Departamento de Ciência da Computação - UFMG

Belo Horizonte, 27 de Julho de 2020.

Dedico esta dissertação a todos que me apoiaram, acreditaram em mim e me incentivaram a aprender e melhorar todos os dias.

Agradecimentos

Primeiramente, gostaria de agradecer aos meus pais: José Adão e Luciana Karem, pois nunca mediram esforços para que eu pudesse ter sempre um ensino de qualidade, me educaram da melhor forma possível e moldaram meu caráter como pessoa. Sempre farei o meu melhor para dar orgulho a vocês. Gostaria de agradecer também à minha irmã, que me traz alegria e amor, além de me incentivar a ler mais. À minha namorada, Caroline, por ter me incentivado e me apoiado em todos os momentos. Amo vocês, família!

Gostaria de agradecer também ao meu orientador Clodoveu, que tornou todo este processo mais interessante e divertido. Sempre foi uma referência para todos no mundo dos dados geográficos, mas agora também é uma referência para mim como pessoa, pela forma de agir e tratar os que estão ao seu redor.

Um agradecimento especial à minha empresa: Digicade, pela oportunidade de aprender e vivenciar na prática problemas diretamente relacionados com minha área de pesquisa. Em particular, gostaria de agradecer ao meu mentor e amigo João Tavares, pelo apoio e companheirismo de sempre. Valeu, pé torto!

Por fim, gostaria de agradecer aos amigos que sempre me apoiaram nessa jornada e aproveitar para me desculpar nos momentos em que me ausentei. Agradeço por sempre me convidarem para sair e por não terem desistido de mim. Alguns destes amigos fizeram parte dessa jornada de aprendizado no ambiente acadêmico, outros, com quem já não tenho tanto contato, me apresentaram ao universo da Ciência da Computação e serei eternamente grato por isto. Em especial: Pablo Luiz, Gabriel Bressani, Gabriel Caixeta, Daniel Fernandes e Douglas Fernandes.

“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda pensou sobre aquilo que todo mundo vê.”
(Arthur Schopenhauer)

Resumo

Arborização Urbana é um tema de grande importância para a sociedade nos dias atuais por todos os benefícios que ela proporciona aos cidadãos, principalmente nos grandes centros. As árvores desempenham um papel muito importante na melhoria da qualidade de vida da população e do meio ambiente, de modo que a arborização vem sendo considerada um dos elementos da infraestrutura urbana. Apesar dessa importância, eventos de queda ou corte preventivo de árvores em áreas públicas têm se tornado frequentes, causando danos materiais e gerando transtornos ao cotidiano das cidades. Partindo de um conjunto expressivo de dados individuais sobre árvores em Belo Horizonte (MG), este projeto busca contribuir para a identificação e solução de problemas de arborização, com especial ênfase na queda de árvores. Técnicas de mineração de dados foram utilizadas para procurar padrões consistentes, expressos como regras de associação ou sequências temporais, que se relacionem à queda de árvores. As análises mostram um conjunto de atributos que aparecem frequentemente junto à árvores que sofreram um processo de queda, e também indicam que cerca de 12% de um conjunto de árvores analisadas na região central de Belo Horizonte possuem características semelhantes às do conjunto de árvores perdidas analisadas. Também foram identificados os componentes informacionais que serviram de base para a construção de uma ferramenta para a coleta de informações geográficas sobre a arborização urbana, visando manter atualizado o banco de dados de arborização com a colaboração de cidadãos.

Palavras-chave: Computação - Teses, Arborização das cidades, Mineração de dados (computação), Banco de dados geográficos

Abstract

Tree coverage in urban spaces is a theme of great importance for current societies, given all the benefits that green spaces provide to the population, especially in large cities. Trees fulfill a very important role to ensure quality of urban living and urban environmental quality, and as a result trees are considered to be an element of urban infrastructure. In spite of the recognition of the importance of tree coverage, events in which a street tree falls or needs to be preventively cut down are quite frequent, damaging property and causing disturbances in the routine of the population. From a rich dataset on urban trees for the city of Belo Horizonte (MG, Brazil), this paper proposes contributions towards the identification and solution of problems related to tree coverage, with special emphasis on felled trees. Data mining techniques are employed in search of consistent patterns, expressed as association rules or temporal sequences, that are related to felling events. The analyses show a set of attributes that frequently appear next to trees that have suffered a fall process, and also indicate that about 12 % of a set of trees analyzed in the central region of Belo Horizonte have similar characteristics to the set of perished trees analyzed. The informational components that served as a basis for building a tool for collecting geographic information on urban afforestation were also identified, aiming to keep the tree database up to date with the collaboration of citizens.

Keywords: Urban Trees, Data Mining, Volunteered Geographic Information

List of Figures

2.1	Etapas da Mineração de Dados. Fonte: adaptado de [18]	19
3.1	Metodologia	23
4.1	Quantidade de árvores por espécie	28
4.2	Renda Média Per Capita	31
4.3	Índice de Desenvolvimento Humano Municipal	32
4.4	Índice de Escolaridade da População Adulta	32
4.5	Planta Geral da Cidade de Minas	33
4.6	Relação entre suporte relativo e confiança	35
4.7	Relação entre <i>lift</i> e <i>leverage</i>	36
4.8	(a) Árvores dentro da Av. Contorno; (b) Árvores na região da Praça da Liberdade	38
4.9	Árvores cluster 1	39
4.10	Mapa de calor das árvores com risco de queda	39
5.1	Fluxograma da aplicação	42
5.2	Diagrama de componentes	44
5.3	Diagrama de classes simplificado	46
5.4	Autenticação e menu lateral	48
5.5	Mapa e atributos das árvores	49
5.6	Cadastro e listagem de ocorrências	49
5.7	(a) identificação de bifurcação; (b) identificação de galhos epicórmicos	50
5.8	(a) identificação de raízes superficiais; (b) indicação de contato com a rede elétrica	50
5.9	Verificação por pares	51

List of Tables

3.1	Atributos utilizados para a aplicação de técnicas de mineração de dados	25
4.1	Árvores por região em Belo Horizonte	29
4.2	Exemplo dos Indicadores Socioeconômicos	30
4.3	UDHs com maior renda média per capita	30
4.4	UDHs com maior índice de desenvolvimento humano municipal	31
4.5	UDHs com maior índice de escolaridade	31
4.6	Regras de Associação	36
4.7	$K = 2$	38
5.1	Métodos implementados na API	47

Contents

1	Introdução	12
1.1	Contextualização e Motivação	12
1.2	Objetivos	14
1.3	Objetivos específicos	14
1.4	Organização do trabalho	15
2	Revisão da literatura	16
2.1	Arborização Urbana	16
2.2	Conceitos de Mineração de Dados	18
2.3	Contribuições voluntárias de dados geográficos	21
2.4	Considerações	22
3	Metodologia	23
4	Um estudo de caso sobre a arborização urbana em Belo Horizonte	27
4.1	Base de Dados	27
4.2	Indicadores Socioeconômicos	29
4.3	Primeiras análises	33
5	VGI	41
5.1	Projeto	41
5.2	Implementação	44
5.3	Funcionamento	47
6	Conclusões	53
	Bibliography	56
A	Atributos na base de dados do SIA/BH	60

Chapter 1

Introdução

1.1 Contextualização e Motivação

Arborização Urbana é um termo que vem sendo utilizado com muita frequência recentemente, visto que árvores em espaços urbanos podem melhorar a qualidade de vida da população e o meio ambiente. Entre os benefícios podemos citar: bem-estar psicológico, efeito estético, sombra para os pedestres e veículos, proteção contra o vento, diminuição da poluição sonora, redução do impacto da água de chuva, auxílio na diminuição da temperatura e preservação da fauna silvestre.

Os estudos de [30, 16] mostram uma ampla análise da literatura e demonstram os benefícios proporcionados pelas árvores no ambiente urbano, as percepções de árvores de rua transmitidas pelos residentes urbanos, que por vezes consideram a arborização um tipo de infraestrutura urbana.

Encaixar as árvores dentro da disposição atual de cidades é uma tarefa complicada, mas é preciso estudá-la. A busca da harmonia entre o crescimento urbano e o meio ambiente é um desafio diário para os órgãos públicos responsáveis por tais setores. Assim como qualquer setor urbano, a arborização das ruas e avenidas está condicionada à qualidade e dedicação ao planejamento.

Na tentativa de dar sua contribuição ao meio ambiente ou de embelezar a frente de suas casas e adjacências, algumas pessoas ultrapassam os limites de sua propriedade e plantam árvores, arbustos e flores no passeio público, canteiros centrais de avenidas, áreas verdes e praças. A ação sem orientação técnica pode, no entanto, causar mais prejuízos que benefícios. O plantio de mudas inapropriadas e em lugares inadequados pode causar acidentes de trânsito ou com a fiação elétrica. Logo, podemos perceber que também é necessário fazer uma análise sobre as condições das árvores que já se encontram nas cidades, para identificar aquelas que podem causar problemas futuros.

Além do plantio de espécies inadequadas podemos citar uma série de outros problemas e desafios da arborização urbana. A falta de espaço para o crescimento adequado das árvores é um exemplo, prejudicando a infiltração da água e de nutrientes nas raízes. Outro

obstáculo são os danos que as espécies sofrem ao longo do tempo. Além disso, colisões de carros, poluição e presença de fiações elétricas são alguns exemplos que prejudicam as árvores no espaço urbano.

Existem algumas iniciativas, como inventários arbóreos municipais, que visam obter e possibilitar análises para compreender a relação entre as árvores e o local onde elas estão inseridas: a compatibilidade entre seu porte (raízes, tronco e copa) e o espaço disponível, as condições sanitárias existentes e sua adaptação. Além dos inventários, alguns municípios possuem manuais técnicos de arborização urbana, que têm como objetivo auxiliar na escolha da espécie, na observação dos elementos no entorno da árvore a ser plantada e do local de plantio de árvores em passeios de vias públicas, áreas livres públicas ou até mesmo em áreas internas de lotes públicos ou privados. São Paulo é um exemplo de município que possui um manual técnico de arborização urbana, conforme mostrado em [36].

No final do mês de maio de 2012, o Instituto Brasileiro de Geografia e Estatística (IBGE) divulgou dados sobre o mais recente Censo Demográfico (2010), que fez a análise de características urbanísticas presentes no entorno dos domicílios brasileiros. Um desses indicadores referiu-se à presença, ou não, de árvores em frente ou próximas a cada uma das residências pesquisadas, conforme mostrado em [27]. Os resultados mostram que existe um déficit de no mínimo 15 milhões de árvores no país e que aproximadamente 32% das residências (quase 50 milhões de pessoas) não são atendidas pelos benefícios trazidos pela arborização urbana. A pesquisa também revela que um terço das cidades brasileiras com mais de 1 milhão de habitantes mantém entre 60% e 77,6% das suas respectivas populações sem o benefício da arborização urbana nas proximidades de sua residência.

Além da escassez de árvores nos grandes centros, outro grande problema são as quedas de árvore. Em Belo Horizonte (Brasil), foram investidos cerca de R\$ 8 milhões em 2018 em ações para diminuir o número de quedas de árvores. De janeiro a julho de 2018 foram executadas 22.437 podas e 4.451 supressões. Em todo o ano de 2017, foram 16.445 podas de árvores e 3.041 supressões¹.

Como a queda de árvores é uma situação recorrente em Belo Horizonte (de janeiro a março de 2018, o Corpo de Bombeiros precisou retirar emergencialmente 516 árvores que caíram na capital e na região metropolitana²), existem várias iniciativas da prefeitura para a retirada de árvores que correm o risco de cair, conforme mostrado em [35].

Mesmo com a existência de planos e diretrizes por parte dos órgãos públicos, é difícil ter acesso à base de dados com as informações sobre as árvores de um município, informações sobre podas realizadas e supressões, pois não é prática corrente a publicação de dados vistos como operacionais. Além disso, as bases de dados existentes, em geral, são

¹https://www.em.com.br/app/noticia/gerais/2018/08/17/interna_gerais,980874/supressoes-de-arvores-em-bh-ja-superam-todo-o-volume-do-ano-passado.shtml

²<https://g1.globo.com/mg/minas-gerais/noticia/prefeitura-define-criterios-para-retirada-de-arvores-que-correm-risco-de-queda-em-bh.ghtml>

desatualizadas e desnormalizadas. A fim de desenvolver um levantamento mais preciso de dados, uma estratégia pode ser utilizar especialistas no assunto para fazer a coleta destas informações, como engenheiros ambientais e florestais, porém este processo pode ser bastante custoso, demorado e difícil de ser mantido pelos órgãos públicos. Para isso, existem estratégias que utilizam a contribuição voluntária de informações para fazer a coleta de dados que não necessitem de especialistas para tal, facilitando assim a manutenção da base de dados e deixando todo o processo menos oneroso. Logo, é possível combinar estas duas abordagens para a manter uma base de dados atualizada sobre a arborização urbana nos grande centros.

A existência de um banco de dados unificado e atualizado pode permitir análises que possam orientar as pessoas e instituições responsáveis pelo manejo das árvores no espaço urbano, além de permitir que a população tenha acesso e conheça parte da flora que está ao seu redor diariamente. Estas análises, inicialmente, têm como objetivo prever queda de árvores, analisando registros e árvores que caíram e buscando padrões; detectar interferências de árvores com outros elementos de infraestrutura; analisar a distribuição de árvores no perímetro urbano; e identificar espécies problemáticas, localizando a ocorrência de espécimes no espaço urbano.

1.2 Objetivos

Observando este cenário, este trabalho tem como objetivo conduzir um estudo sobre a arborização urbana, utilizando um banco de dados geográfico existente e técnicas de mineração de dados para procurar padrões consistentes, como regras de associação e sequências temporais, para encontrar padrões desconhecidos e detectar relacionamentos sistemáticos entre atributos das árvores, produzindo assim novos subconjuntos de dados que forneçam indicações para solucionar problemas relacionados com o tema, como a queda de árvores.

1.3 Objetivos específicos

A seguir são exibidos os objetivos específicos deste trabalho:

- Obter e tratar dados geográficos de arborização urbana de Belo Horizonte;

- Integrar os dados de arborização a outros dados do espaço urbano, de modo a permitir a análise comparativa e formulação de hipóteses para padrões;
- Analisar dados usando regras de associação e agrupamento como técnicas de mineração de dados, de modo a buscar padrões relacionados com a queda de árvores; e
- Projetar recursos computacionais que permitam que cidadãos comuns contribuam com a manutenção de dados de arborização, de modo a aumentar a qualidade de futuras análises e planos.

1.4 Organização do trabalho

O restante desta dissertação está organizada como segue. O Capítulo 2 discute uma série de trabalhos relacionados que se preocupam com a arborização urbana e trabalhos que utilizam mineração de dados para fazer análises em grandes conjuntos de dados, enquanto o Capítulo 3 apresenta os dados que são utilizados para o desenvolvimento do trabalho aqui proposto e a metodologia proposta. No Capítulo 4 é mostrado um estudo de caso utilizando dados de uma base de dados sobre arborização urbana de uma cidade brasileira. O Capítulo 5 traz uma ferramenta de coleta de dados por voluntários que foi desenvolvida com o objetivo de auxiliar no processo de coleta e atualização dos dados relacionados com a arborização urbana. Por fim, o Capítulo 6 traz conclusões e discute trabalhos futuros.

Chapter 2

Revisão da literatura

Existe uma variedade de esforços na literatura que propõem estudos envolvendo arborização urbana. [33] mostra que várias condições são exigidas de uma árvore para que possa ser utilizada sem acarretar inconvenientes como, por exemplo resistência a pragas e doenças, troncos e ramos resistentes (para evitar queda), raízes profundas, entre outras.

Diante da quantidade de estudos já realizados, neste capítulo são apresentados os principais trabalhos que serviram como inspiração para o desenvolvimento da pesquisa aqui realizada. Este capítulo é, portanto, dividido em quatro seções, sendo que na Seção 2.1 são apresentados os trabalhos relacionados com o tema de arborização urbana. Na Seção 2.2 são apresentados os trabalhos que utilizam mineração de dados para tratar dados similares aos que serão utilizados nesta dissertação. A Seção 2.3 traz trabalhos relacionados com a área de VGI (*Volunteered Geographic Information*), a fim de buscar na literatura as melhores formas de desenvolver ferramentas que utilizam a estratégia de envolver grandes grupos de cidadãos em iniciativas de coleta e manutenção de dados espaciais. Por fim, na seção 2.4 são discutidas as principais diferenças entre o que já se encontra disponível na literatura e a pesquisa realizada nesta dissertação.

2.1 Arborização Urbana

Existe uma grande quantidade de estudos relacionados à arborização urbana na área da Biologia e Ciências da Terra. O estudo de [9] mostra a falta de organização das cidades quanto ao planejamento e manejo das árvores em vias urbanas e também ressalta a importância e os benefícios do verde urbano. Um inventário de árvores utiliza fundamentos de amostragem para estimar os parâmetros quantitativos, como volume e área basal, assim como os qualitativos, como qualidade de fustes e valoração de espécies. A criação de inventários é uma prática comum em algumas cidades para se ter um maior controle sobre as espécies presentes em cada município. [14] e [11] são exemplos de inventários de

árvores de cidades brasileiras.

Em Chicago (EUA), [38] estimaram que o plantio de 50 mil árvores e sua manutenção por 30 anos custaria US\$ 8,4 milhões, enquanto os benefícios conferidos pelas árvores deveriam ser de US\$ 23,5 milhões. Na cidade de Modesto (EUA), com mais de 90 mil árvores públicas (uma árvore para cada dois residentes), foi realizado estudo para saber se os benefícios da arborização urbana justificavam o orçamento anual de US\$ 2 milhões com sua manutenção. Os resultados indicaram que os benefícios obtidos excediam os custos de manutenção de árvores por um fator de aproximadamente 2.

Companhias de energia elétrica são diretamente impactadas pela arborização urbana, visto que precisam se preocupar com a relação entre as árvores e a rede elétrica aérea para evitar acidentes e desastres. Por isso, em algumas cidades as companhias elétricas dividem com a prefeitura a responsabilidade pela poda das árvores.

[8] divide as práticas de manutenção de árvores em:

- medida preventiva – evita e previne eventuais problemas que as árvores possam sofrer nas ruas ou, ainda, a superação de algum dano com pouca significância;
- medida remediadora – atenua uma falta ou um mal e pode reparar ou corrigir um problema ocorrido com a árvore no meio urbano, geralmente ao tronco, ocorridos por fatores naturais ou injúrias mecânicas, tais como: acidentes com veículos, ventos e vandalismo; e
- medida supressória – se destina a suprimir, alterar ou eliminar a árvore do local por fatores relativos ao próprio indivíduo ou o meio urbano. É aplicável em caso de árvores com doenças ou pragas, risco de queda ou morte comprovada; a remoção de flores e frutos desagradáveis ou com princípios alérgicos ou ainda a remoção de árvores a pedido da população.

No cenário brasileiro, e especificamente em companhias elétricas, dentre os processos de O&M (Organização e Métodos), destaca-se o Manejo de Vegetação, composto pelo planejamento e execução da poda de árvores do ambiente urbano e da limpeza da faixa de servidão das redes e linhas rurais. Nesse processo, são aplicados cerca de 80% dos recursos de despesa operacional da manutenção preventiva, com reflexos ainda em corretiva nos casos de ocorrências, além de representar a maior causa de interrupções sustentadas do fornecimento de energia¹. Adicionalmente, temos neste processo, um significativo risco de impacto ambiental por ter intervenção direta no meio ambiente, cujos procedimentos de execução estão diretamente relacionados a requisitos legais que extrapolam as instruções internas da companhia. Portanto, é notório que a busca contínua pelo aprimoramento do processo de manejo da arborização tem efeito direto e imediato no resultado operacional em termos de custos e qualidade da energia, bem como na gestão ambiental, minimizando

¹http://www.cemig.com.br/pt-br/a_cemig/nossos_negocios/Paginas/indicadores_de_qualidade.aspx

os riscos de aplicação de sanções legais e pecuniárias referentes à legislação ambiental aplicável.

A Companhia Energética de Minas Gerais (CEMIG) é uma das principais concessionárias de energia elétrica do Brasil. Segundo o seu Manual de Arborização, exibido em [12], quando for preciso realizar uma intervenção em uma área de arborização urbana, é importante planejá-la, para evitar ou minimizar os danos sobre ela mesma e sobre os outros seres vivos que com ela interagem, incluindo os seres humanos. Atualmente existem alguns critérios estabelecidos para definir se uma árvore possui risco de queda ou não, que são: vigor da árvore, elevação do solo, presença de pragas e deformação do tronco. [33] mostra que várias condições são exigidas de uma espécie de árvore para que possa ser utilizada nas ruas sem acarretar inconvenientes como, por exemplo resistência a pragas e doenças, troncos e ramos resistentes (para evitar queda), raízes profundas, entre outras.

[24] destacam que ao considerar as árvores como equiparadas a outros elementos de infraestrutura urbana, pode-se valorizá-las em importância e em valor monetário, de forma equivalente aos demais. [3] buscam apresentar os benefícios das árvores urbanas como entidades mensuráveis, estimando de forma unitária o valor de cada árvore. Em [31] é calculado o valor da arborização urbana em Oakland, nos Estados Unidos, através da fórmula do valor da árvore (*tree-value formula*), para mostrar que entender o valor das árvores pode dar a órgãos públicos uma base melhor para gestão de árvores urbanas. Por sua vez, [17] fazem uma comparação entre os benefícios e os custos do verde urbano, estes, muitas vezes, maiores do que o necessário por causa do planejamento mal feito.

Atualmente existem poucas iniciativas para proteção das áreas de arborização urbana, mas isto não quer dizer que o tema não seja de extrema importância para toda a população, demandando um trabalho de conscientização para aumentar a preocupação das pessoas com tal tema. Os benefícios são indiscutíveis e a justificativa para se fazer um plano de manejo para estes espaços verdes vai ao encontro a valores culturais, históricos, familiares, ecológicos, ou mesmo por motivos de exigência legal e de valorização financeira.

2.2 Conceitos de Mineração de Dados

A Mineração de Dados é uma linha de pesquisa pertencente ao campo da Ciência da Computação que tem por objetivo oferecer estratégias para extrair informações que estejam implícitas em grandes bases de dados. Em geral, busca-se encontrar informações previamente desconhecidas e potencialmente úteis para fornecer *insights* para tomadas de decisões das instituições como, por exemplo ajudar a detectar tendências de vendas, desenvolver campanhas de marketing mais inteligentes e prever com precisão a fidelidade

do cliente.

[41] fornecem uma visão geral ampla, porém aprofundada, da mineração de dados, integrando conceitos de aprendizado de máquina e estatística. O livro inclui alguns exemplos de análise de dados, mineração de padrões, agrupamento e classificação e foi a principal literatura utilizada como referência para o desenvolvimento das etapas de mineração de dados desta dissertação.

As principais etapas de um processo de Mineração de Dados podem ser divididas em três etapas principais: (1) Tratamento e organização do banco de dados; (2) Mineração de Dados; e (3) Avaliação e visualização. A figura 2.1 mostra todas as etapas envolvidas em um fluxo para a mineração de dados.

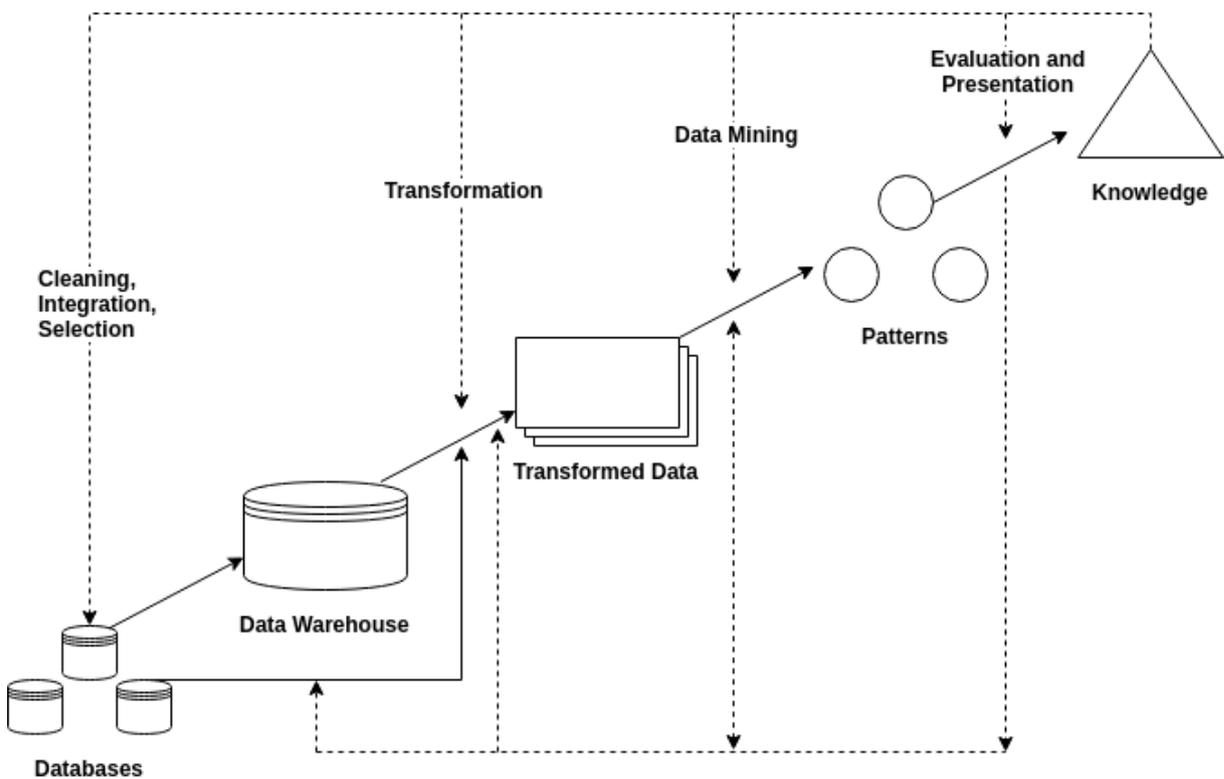


Figure 2.1: Etapas da Mineração de Dados. Fonte: adaptado de [18]

2.2.1 Tratamento e organização do banco de dados

Esta etapa tem como objetivo o tratamento e organização do banco de dados inicial que será utilizado nos primeiros experimentos. Diante disso, as tarefas compreendidas por esta etapa são: tratamento e organização.

- Tratamento dos dados: Durante esta etapa são removidos dados inválidos, seja por não apresentarem os atributos obrigatórios para realização de futuras análises ou por possuírem caracteres inválidos, por exemplo, possuir caracteres não-numéricos em atributos que só podem possuir dados numéricos; e
- Organização: Após o tratamento, os dados são organizados de forma a facilitar a fase de execução dos experimentos. Além disso, são realizados pré-processamentos com o objetivo de extrair novas características. A etapa de organização dos dados é fundamental para possibilitar uma melhor visualização dos dados, visto que o dado bruto é, em geral, muito desorganizado.

2.2.2 Mineração de Dados

A área de estudo da Mineração de Dados consiste no uso de um conjunto técnicas que permitem automatizar a busca em grandes volumes de dados por padrões e tendências que não são detectáveis por análises mais simples. Contudo, antes de aplicar as técnicas é necessário fazer um tratamento e organização da base de dados, este processo foi explicado na seção anterior.

Os algoritmos fundamentais em mineração de dados e a análise dos resultados obtidos por eles formam a base para o emergente campo da Ciência de Dados, que inclui métodos automatizados para analisar padrões e modelos para todos os tipos de dados, com aplicações que vão desde descobertas científicas a *business intelligence* e *analytics*, como mostrado em [41].

As várias técnicas de mineração de dados existentes permitem segmentar os dados e avaliar a probabilidade de ocorrência de determinados eventos no futuro. As próximas subseções explicam de forma sucinta as principais abordagens da etapa de mineração de dados.

Regras de Associação

Regras de associação representam padrões de relacionamento entre itens de uma base de dados. Uma de suas típicas aplicações é a análise de transações de compras (*market basket analysis*), um processo que examina padrões de compras de consumidores para determinar produtos que costumam ser adquiridos em conjunto.

Em [1, 2] é mostrado como a análise de padrões frequentes e regras de associação podem ser valiosas ao analisar grandes bases de dados, sendo possível entender melhor

os dados em estudo, analisando os atributos que aparecem em conjunto em uma certa frequência.

Agrupamento

Alguns algoritmos e estratégias de agrupamento são mostrados em [7], esta técnica de mineração de dados visa condensar dados com atributos similares em um mesmo grupo e pode caracterizar grupos de dados, possibilitando análises e ações específicas para cada um deles.

2.3 Contribuições voluntárias de dados geográficos

Os termos *Crowdsourcing* e *Volunteered Geographic Information* (VGI) surgiram nos últimos anos para caracterizar um conjunto de métodos e técnicas voltadas a permitir que cidadãos voluntários constituam uma importante fonte de informações que podem auxiliar no desenvolvimento de ferramentas para diversos fins. No caso de VGI, cidadãos comuns usam ferramentas para este propósito de modo a contribuir com suas visões sobre dados geográficos na web, conforme mostrado por [22, 21]. [19] levantam alguns aspectos importantes sobre a credibilidade das informações geográficas coletadas voluntariamente, mostrando como articular estratégias para discernir a confiabilidade do dado fornecido e também os motivos que levam as pessoas a contribuírem com tais ferramentas. [37], por sua vez, sugerem novas medidas e indicadores de qualidade para alguns tipos de VGI e também mostram os métodos de avaliação de qualidade dos métodos existentes.

Um aspecto importante a ser considerado ao desenvolver uma ferramenta de VGI são as questões éticas e legais. [34] explica que as atividades relacionadas a VGI levantam uma variedade de questões legais, desde a propriedade intelectual à responsabilidade, difamação e privacidade. Essas questões podem afetar os direitos e obrigações de todos aqueles que usam ou participam das atividades da ferramenta.

[26] apresentam uma revisão sistêmica da literatura para avaliar como a VGI vem sendo utilizada como fonte de informação para auxiliar o gerenciamento de desastres naturais, algo bem próximo ao que está sendo proposto aqui nesta dissertação, visto que quedas de árvores também se encaixam na categoria de desastres naturais.

2.4 Considerações

Várias cidades, no Brasil e no mundo, contam com a presença, mesmo que em uma quantidade menor que o ideal, de arborização em suas vias, e que pode ser considerada patrimônio natural, cultural e histórico de valor inestimável para as pessoas que ali vivem. Entretanto, algumas dessas cidades registram também a ocorrência de inúmeros conflitos entre essa arborização e os vários equipamentos urbanos instalados, que estão se tornando cada vez mais comuns e que vêm causando transtornos à população. A constatação deste fato leva à necessidade de se buscar novas informações e ações que ajudem a solucionar problemas relacionados com a arborização urbana como, por exemplo, a queda de árvores.

Podemos perceber que existem vários trabalhos na literatura que abordam a questão da arborização urbana, entretanto, até onde pudemos verificar, não encontramos trabalhos relacionados que usem técnicas de mineração de dados para detectar padrões de problemas ou antecipar o risco de queda de árvores urbanas. Em geral, os estudos encontrados buscam classificar espécies de acordo com imagens de satélite para analisar a cobertura da vegetação urbana [39, 28].

Diante da existência de dados detalhados sobre cada árvore urbana em Belo Horizonte, um conjunto de dados georreferenciados que vem sendo trabalhado na cidade desde a década de 1990, este trabalho busca utilizar a Mineração de Dados e a contribuição de cidadãos voluntários para explorar as possibilidades de análise no sentido de contribuir com o esforço de manutenção da arborização na cidade, e prover um modelo para outras localidades.

Chapter 3

Metodologia

Nos capítulos anteriores mostramos que as árvores melhoram a qualidade de vida em nossas cidades por inúmeras razões. Visando ajudar a repensar sobre a arborização urbana das cidades atuais e as cidades do futuro, este trabalho tem como objetivo criar um modelo que possa ser replicado em inúmeras cidades através da descoberta de conhecimento e coleta de dados relacionados com a distribuição arbórea em centros urbanos. A figura 3.1 mostra o processo de trabalho utilizado nesta dissertação.

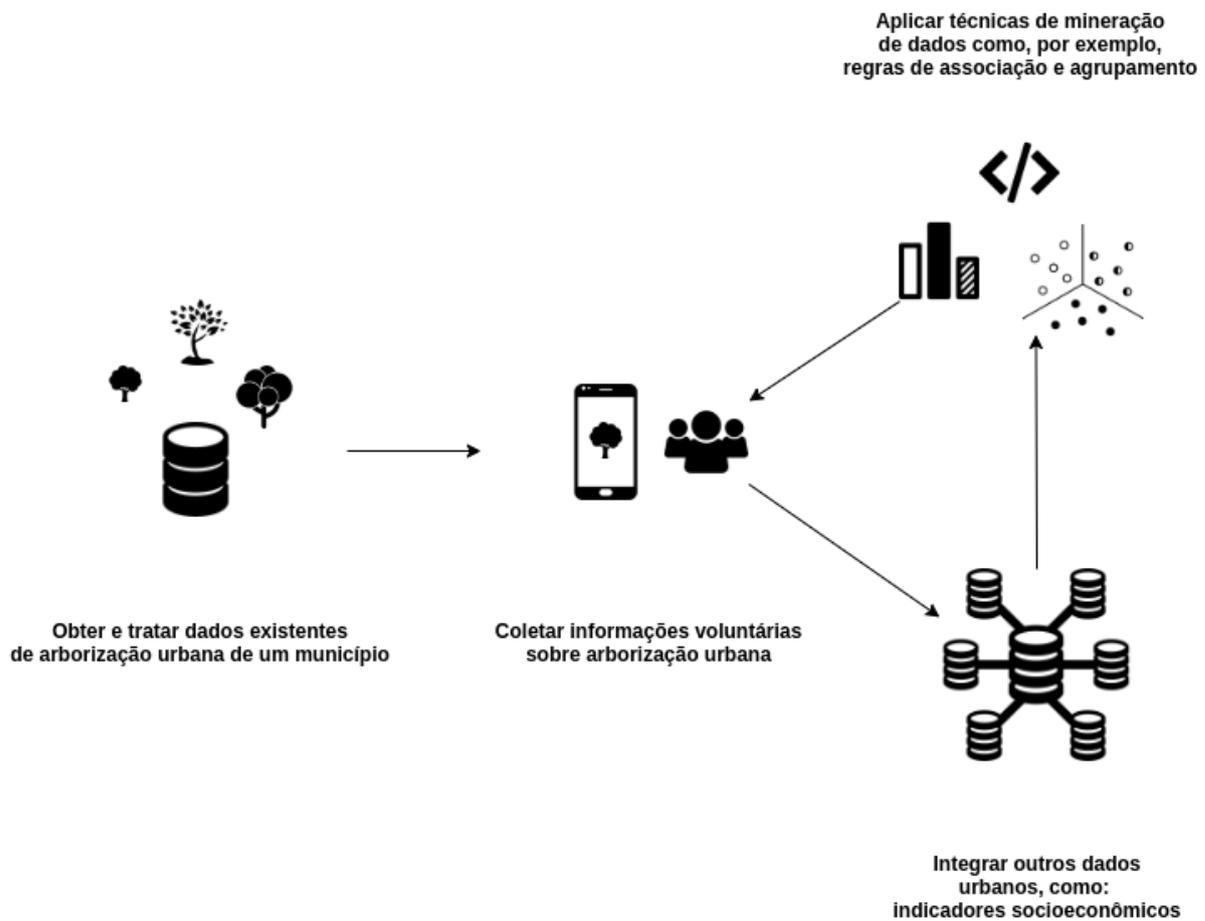


Figure 3.1: Metodologia

O termo *Knowledge Discovery in Databases* (KDD), mostrado em [20], refere-se ao amplo processo de busca de conhecimento em dados e enfatiza a aplicação "de alto

nível” de métodos específicos de mineração de dados. Estes métodos estão relacionados com diversas áreas da computação: aprendizado de máquina, reconhecimento de padrões, bancos de dados, estatística, inteligência artificial e visualização de dados.

O objetivo unificador do processo KDD é extrair conhecimento dos dados no contexto de grandes bases de dados. Isso é feito usando métodos de mineração de dados (algoritmos) para extrair (identificar) o que é considerado conhecimento, de acordo com as especificações de medidas e limites, usando uma base de dados juntamente com qualquer pré-processamento, subamostragem e transformação necessários desse conjunto de dados.

Para que este processo seja possível, é necessário que tenhamos uma base de dados. No cenário de arborização urbana, o ideal seria a existência de um inventário de árvores da cidade em que se deseja fazer o estudo. [14] e [11] são exemplos de inventários de árvores de cidades brasileiras e que podem ser utilizados como referência para a criação de um inventário em uma cidade que ainda não possua.

A Tabela 3.1 mostra alguns dos atributos de uma base existente de dados geográficos sobre arborização urbana que foram utilizados nas análises presentes neste trabalho e suas respectivas siglas. Destacamos a informação sobre brotação e galhos epicórmicos, ou seja, gemas adventícias do tronco que apresentam uma ligação deficiente com a parte central do tronco, podendo se constituir em fator de risco no futuro.

Após extrair o conhecimento dos dados é possível fazer análises mais aprofundadas, sendo possível confirmar ou refutar as hipóteses antes levantadas. Entretanto, a forma de exibir os resultados gerados pelo processo de KDD é de suma importância para extrair as respostas buscadas, como mostrado em [25]. A visualização de dados é o ato de pegar informações (dados) e colocá-las em um contexto visual, como um mapa ou gráfico. As visualizações de dados facilitam o entendimento de conjunto de dados grandes ou pequenos, e a visualização também facilita a detecção de padrões, tendências e discrepâncias em grupos de dados. Boas visualizações de dados devem colocar significado em conjuntos de dados complicados, para que suas mensagens sejam claras e concisas.

As técnicas de mineração de dados escolhidas são: Associação e Agrupamento. O uso da técnica de regras de associação tem como objetivo avaliar se existe alguma regra de associação entre árvores perecidas e outros atributos, a fim de identificar possíveis causas de quedas de árvores. Enquanto o uso da técnica de agrupamento tem como objetivo verificar se existem grupos que consigam distinguir árvores que caíram e árvores que possuem grandes chances de queda das demais árvores. Como o número de atributos é muito grande, é necessário escolher aqueles que são os mais relevantes para identificar árvores que caíram. Atualmente existem alguns critérios estabelecidos para definir se uma árvore possui risco de queda ou não, que são: vigor da árvore, elevação do solo, presença de pragas e deformação do tronco. Como alguns desses atributos são subjetivos, como o vigor da árvore, foi escolhido outro conjunto de atributos na tentativa de reduzir a

Sigla	Atributo	Observação
PERECIDA	Perecida	Árvore que sofreu queda
PODA_NECESSARIA	Poda Necessária	Árvore que necessita de algum processo de poda
PODADA	Podada	Árvore que sofreu processo de poda recentemente
BIFURCADA	Bifurcada	Tipo de bifurcação da árvore
ENTOUCEIRADA	Entouceirada	Tipo de bifurcação da árvore
BCA	Base com aterro	-
BCBE	Base com brotação epicórmica	-
BCF	Base com fungo	-
BCPI	Base com presença de insetos	-
CCEP	Copa com erva de passarinho	-
CCFR	Copa com folhagem rala	-
CCF	Copa com fungos	-
CCGO	Copa com galhos ocos	-
CCGS	Copa com galhos secos	-
CCSPR	Copa com sinal de poda rebaixada	-
CCPU	Copa com poda unilateral	-
ESF	Elevação do solo e fissuras	-
PGC	Presença de galhos com casca	-
CCGE	Copa com galhos epicórmicos	-
PERFILHAMENTO	Perfilhamento	-
PRA	Presença de raízes adventícias	-
RCE	Raízes cortadas ou estrangulante	-
EC	Relação entre copa e cabos	-
S2M	Semáforo a menos de 2 metros	-
AR	Superficialidade da raiz	-
TCE	Tronco com corpos estranhos	-
TCF	Tronco com fungo	-
TCG	Tronco com grade	-
TCI	Tronco com inseto	-

Table 3.1: Atributos utilizados para a aplicação de técnicas de mineração de dados

dimensionalidade dos dados e encontrar um novo subconjunto de dados que represente bem as árvores perecidas.

Os atributos escolhidos para a etapa de agrupamento foram: árvore bifurcada ou não, presença de galhos epicórmicos, presença de rede elétrica ou ramal que encoste na árvore, e relação entre a copa e cabos. Esses atributos foram escolhidos com base nas descobertas obtidas ao aplicar a técnica de regras de associação, visto que os resultados mostraram que, em geral, árvores perecidas possuem copa com galhos epicórmicos e isso pode ser decorrente de podas malfeitas. Além disso, árvores que possuem relação direta com a rede elétrica estão mais sujeitas a passarem por podas em intervalos menores do que as demais, para evitar riscos maiores e prejuízos à população, e principalmente, às companhias elétricas. Estes atributos foram transformados em variáveis categóricas antes de serem passados com parâmetro para o algoritmo.

Além das aplicações das técnicas de mineração de dados mencionadas, também são exibidas visualizações que permitem visualizar a distribuição espacial dos dados. Os resultados obtidos após esta etapa são utilizados como ponto de partida para o desenvolvimento de uma nova ferramenta para a coleta de dados de forma voluntária, ou seja, VGI. Desta forma, é possível definir, antes do início do desenvolvimento, quais os atributos são

os mais importantes para serem coletados, fazendo com que a aplicação colete apenas o necessário para alimentar os modelos de mineração desenvolvidos e modelos futuros. Além disso, a redução do número de atributos coletados irá melhorar a usabilidade da ferramenta, tornando a ferramenta fácil de usar e não cansando os usuários.

Utilizando o processo de descoberta de conhecimento, juntamente com um inventário de árvores existentes, percebemos que a atualização desses dados é difícil, visto que este processo pode ser longo, de acordo com o tamanho da área urbana a ser analisada, e custoso, caso a coleta dos dados seja feita por especialistas (engenheiros florestais e ambientais, por exemplo).

O uso de estratégias como *crowdsourcing* ou *volunteered geographic information* (VGI) possibilitam que a base de dados existente seja atualizada por cidadãos comuns que possuam algum engajamento com o tema tratado, [22, 21]. Desta forma, é possível cobrir uma área maior, com menos recurso e de maneira mais rápida. Porém, é preciso adotar métodos para validar as informações que estão sendo inseridas e definir um processo de aprovação, para que o dado cadastrado passe a ser um registro oficial na base de dados.

Portanto, com uma base de dados unificada e atualizada é possível repetir análises feitas na etapa de extração de conhecimento dos dados, refutar ou comprovar os resultados e gerar novas análises com os dados atualizados.

No cenário da arborização urbana, a etapa de VGI permitiria também uma série de análises temporais, visto que a atualização dos dados das árvores não seria feita apenas em períodos sazonais, como ocorre nos casos dos inventários arbóreos. Além disso, é um tema que traz muito engajamento por parte da população, pelos benefícios sociais e ambientais trazidos pela arborização urbana ou pelos transtornos causados por problemas associados com quedas de árvores no ambiente urbano.

O próximo capítulo mostra a aplicação da metodologia aqui exposta em um estudo de caso utilizando dados de arborização urbana da cidade de Belo Horizonte, Minas Gerais. Dessa forma, será explanado o processo de aplicação dessa metodologia em um cenário real, servindo de base para aplicação em outras cidades futuramente.

Chapter 4

Um estudo de caso sobre a arborização urbana em Belo Horizonte

Este estudo de caso visa mostrar uma análise dos dados sobre arborização urbana da cidade de Belo Horizonte, Minas Gerais. Uma síntese deste capítulo foi publicada na edição de 2019 do GeoInfo (Brazilian Symposium on Geoinformatics)¹ com o título *Analyzing Data On The Tree Coverage Of A Large City*, [32].

As próximas seções estão divididas como se segue. A seção 4.1 descreve os dados e as várias fontes que foram utilizadas para obtê-los. Por sua vez, a seção 4.2 exibe uma análise dos dados obtidos com indicadores socioeconômicos da cidade de Belo Horizonte. Por fim, a seção 4.3 exibe o resultado da aplicação de algumas técnicas de mineração de dados em uma amostra do conjunto de dados obtidos.

4.1 Base de Dados

Os dados de arborização são provenientes do Sistema de Informações e Inventário das Árvores de Belo Horizonte - SIA/BH. O “Inventário das Árvores de BH” resulta de um convênio assinado no início de 2011, entre a Prefeitura de Belo Horizonte, a Cemig e a Universidade Federal de Lavras (UFLA), que teve por objetivo estudar cerca de 300 mil árvores nas ruas da cidade. O banco de dados resultante foi cedido pelo órgão municipal responsável, embora não faça parte dos conjuntos de dados abertos² publicados pela Prefeitura de Belo Horizonte³.

Para cada árvore cadastrada, foram coletados diversos dados, referentes à espécie

¹<http://geoinfo.info/geoinfo2019/>

²<https://dados.pbh.gov.br/>

³<http://bhmap.pbh.gov.br/v2/mapa>

e à situação de plantio e saúde. Os dados foram obtidos por engenheiros florestais e engenheiros ambientais, que foram a campo portando dispositivos móveis configurados com uma aplicação especialista, que permitia também a coleta de fotos para cada cadastro. A aplicação possuía um questionário com 57 atributos relacionados às características, estado de saúde e situação de plantio da árvore, como altura, espécie, diâmetro transversal, relação com rede elétrica ou ramal, relação entre copa e cabos, fungos, entre outros.

O apêndice A mostra os diferentes grupos de atributos e os possíveis valores que eles podem assumir na base de dados.

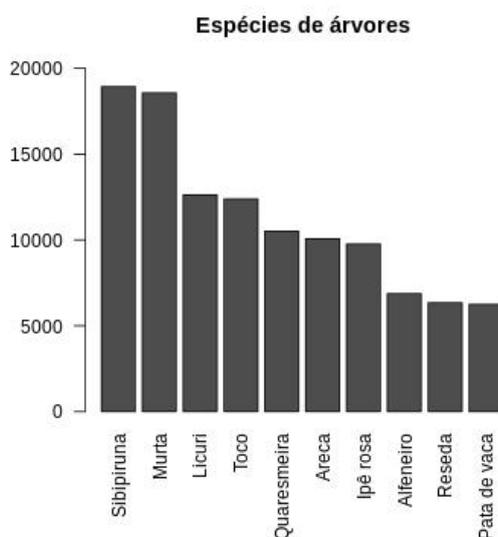


Figure 4.1: Quantidade de árvores por espécie

A Figura 4.1 mostra a quantidade de árvores das espécies mais comuns presentes na base de dados. Na cidade de Belo Horizonte a espécie predominante é a sibipiruna, com mais de 18.000 unidades catalogadas. Essa é uma árvore nativa brasileira, de grande porte e que está presente na maioria dos grandes centros brasileiros. A sibipiruna é bastante encontrada em praças públicas e grandes jardins, porém não deve ser plantada em ruas com calçadas estreitas e com rede elétrica devido ao seu porte. Apesar disso, este tipo de árvore possui grande capacidade de redução de temperatura sob seu dossel (resultado da sobreposição dos galhos e folhas das árvores) e é muito indicada para cidades que apresentam problemas com ilhas de calor.

A Tabela 4.1 mostra que a maior concentração de árvores se encontra na região Centro-Sul de Belo Horizonte, que é a região mais antiga, central e economicamente mais importante da cidade, e que possui um enorme fluxo de pessoas diariamente, com grande quantidade de comércios e residências.

Região	Quantidade de árvores
Centro-Sul	91.792
Pampulha	63.733
Noroeste	59.320
Oeste	53.880
Leste	31.151

Table 4.1: Árvores por região em Belo Horizonte

4.2 Indicadores Socioeconômicos

Para efeito e análise de comparação com os dados de arborização, foram utilizados os dados de índices socioeconômicos do Atlas do Desenvolvimento Humano no Brasil. O Atlas é uma plataforma de consulta ao Índice de Desenvolvimento Humano Municipal (IDHM) de 5.565 municípios brasileiros, 27 Unidades da Federação, 21 Regiões Metropolitanas e 3 Regiões Integradas de Desenvolvimento e suas respectivas Unidades de Desenvolvimento Humano (UDHs). Esta plataforma foi construída por meio de uma iniciativa entre o Programa das Nações Unidas para o Desenvolvimento, o Instituto de Pesquisa Econômica Aplicada e a Fundação João Pinheiro.

Os dados utilizados neste trabalho são das UDHs de Belo Horizonte. De acordo com o website do Atlas do Desenvolvimento Humano no Brasil, as UDHs foram delimitadas buscando gerar áreas mais homogêneas, do ponto de vista das condições socioeconômicas, do que as áreas de ponderação do Instituto Brasileiro de Geografia e Estatística (IBGE). Ou seja, elas são construídas com o objetivo de melhor captar a diversidade de situações relacionadas com o desenvolvimento humano que ocorre no interior dos espaços intrametropolitanos, notadamente em seus grandes municípios, para desvendar o que é escondido pelas médias municipais agregadas. Enquanto a lógica das áreas de ponderação do IBGE atendem a quesitos técnicos relacionados ao processo de coleta e amostragem, as UDHs estão voltadas para a análise espacial das Regiões Metropolitanas por meio de recortes espaciais de maior homogeneidade socioeconômica, com o objetivo de retratar as desigualdades intrametropolitanas de forma mais contundente.⁴

Sendo assim, é possível acessar de forma pública os dados de Belo Horizonte que foram utilizados no presente trabalho ⁵. A pasta que é disponibilizada para download contém um arquivo shape file com a divisão espacial das UDHs em Belo Horizonte, um arquivo csv contendo os dados do censo de 2000 e 2010 e, também, um csv com o dicionário de cada sigla.

O arquivo csv com os dados do censo das UDHs possui mais de 227 atributos,

⁴[5]

⁵<http://www.atlasbrasil.org.br/2013/pt/download>

entre eles estão os atributos que foram utilizados neste trabalho: Renda Média Per Capita (RDPC), Escolaridade Da População Adulta (ESCOLARIDADE) e IDHM. O primeiro indicador, a RDPC é calculado a partir da razão entre o somatório da renda de todos os indivíduos residentes em domicílios particulares permanentes e o número total desses indivíduos, sendo valores em reais de 01/agosto de 2010. O segundo índice utilizado, a escolaridade média, representa o nível de escolaridade da população adulta, ou seja, o percentual de pessoas com 18 anos ou mais que possuem o fundamental completo. O último é o IDHM, calculado por meio de uma média geométrica dos índices de renda, educação e longevidade, com pesos iguais.

A Tabela 4.2 representa um exemplo de como os índices utilizados neste trabalho estão dispostos na base de dados.

NOME UDH	RDPC	ESCOLARIDADE	IDHM
Lindéia / Jatobá	654,54	0,619	0,736
Solar do Barreiro	408,57	0,464	0,653
Bonsucesso / São João	394,43	0,428	0,637
Vila Olhos d'Água	408,57	0,464	0,653
Vila Pilar	408,57	0,464	0,653

Table 4.2: Exemplo dos Indicadores Socioeconômicos

Primeiramente, foi proposta uma análise dos índices socioeconômicos de Belo Horizonte. Observa-se que 68% das UDHS de Belo Horizonte possuem uma renda média per capita abaixo de R\$ 778,00 e apenas 5,8% das UDHS possuem renda média per capita acima de R\$ 2.349,00. Para fazer um comparativo com a região que essas UDHS estão localizadas, na Tabela 4.3 é possível visualizar as cinco UDHS com maior renda média per capita, sendo que todas elas estão localizadas na região centro-sul da cidade.

Renda Média Per Capita	UDH	Região
7516,82	Belvedere	Centro - Sul
7516,82	Mangabeiras / Parque das Mangabeiras	Centro - Sul
6153,01	Santo Agostinho / Lourdes	Centro - Sul
6032,84	Savassi : Boa Viagem / Funcionários	Centro - Sul
5577,09	Cidade Jardim / Santa Lúcia / São Bento	Centro - Sul

Table 4.3: UDHS com maior renda média per capita

Seguindo a mesma proposta acima, as Tabelas 4.4 e 4.5 são referentes ao IDHM e o índice de escolaridade da população adulta, das cinco UDHS com os maiores índices. Assim, foi possível observar que apesar de haver uma alteração entre as UDHS que lideraram os rankings, a região centro-sul permaneceu, sendo a região com os maiores indicadores.

O primeiro mapa da Figura 4.2 representa a renda média per capita das UDHS de Belo Horizonte, quanto mais escuro a tonalidade do vermelho maior o valor da renda média

IDHM	UDH	Região
0,955	Santo Agostinho / Lourdes	Centro - Sul
0,954	Savassi : Boa Viagem / Funcionários	Centro - Sul
0,951	Cruzeiro / Comiteco	Centro - Sul
0,951	Belvedere	Centro - Sul
0,951	Mangabeiras / Parque das Mangabeiras	Centro - Sul

Table 4.4: UDHs com maior índice de desenvolvimento humano municipal

Índice de Escolaridade	UDH	Região
0,949	Santo Antônio / São Pedro	Centro - Sul
0,948	Santo Agostinho / Lourdes	Centro - Sul
0,946	Cidade Jardim / Santa Lúcia / São Bento	Centro - Sul
0,946	Savassi : Boa Viagem / Funcionários	Centro - Sul
0,941	Cruzeiro / Comiteco	Centro - Sul

Table 4.5: UDHs com maior índice de escolaridade

per capita. O segundo mapa mostrado na Figura 4.2 foi construído com a sobreposição de uma mapa de pontos de densidade das árvores e o mapa do indicador de renda média per capita. O mapa de pontos de densidade é representado pela mancha e quanto mais forte a tonalidade de vermelho representa maior quantidade de árvores no local. Nessa imagem conseguimos visualizar a presença maior de árvores nas regiões com maior índice de renda média per capita.

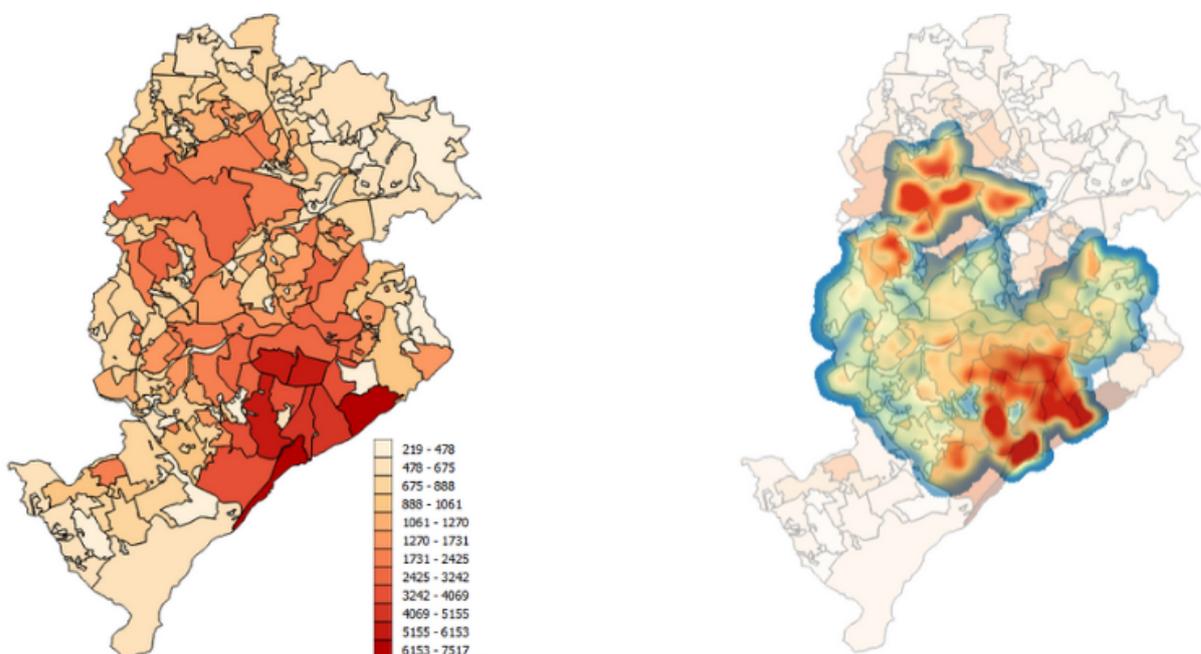


Figure 4.2: Renda Média Per Capita

As figuras mostradas em 4.3 e 4.4 representam o IDHM e o índice de escolaridade da

população adulta, respectivamente, das UDHs de Belo Horizonte. Assim, nos mapas em 4.3 quanto mais escura a tonalidade do azul, maior o índice de desenvolvimento humano municipal. Já nas figuras exibidas em 4.4, quanto mais forte a tonalidade de roxo maior o índice.

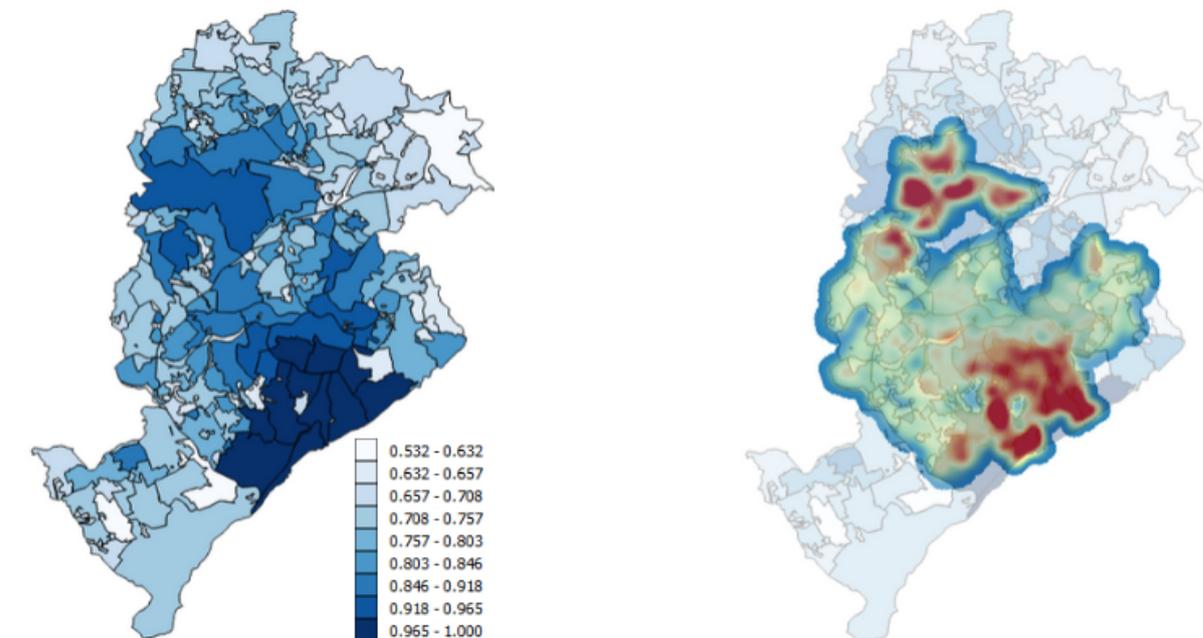


Figure 4.3: Índice de Desenvolvimento Humano Municipal

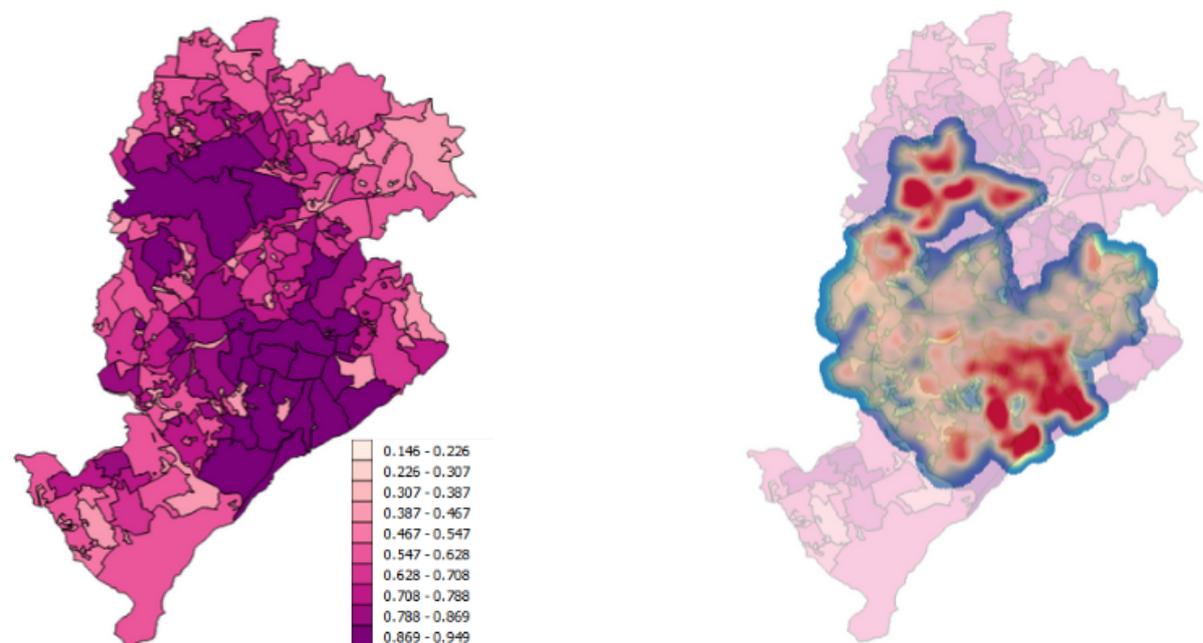


Figure 4.4: Índice de Escolaridade da População Adulta

Desta forma, é possível visualizar que as regiões que possuem maior índice de desenvolvimento humano, possuem também maior arborização urbana, o mesmo também

pode ser observado no caso do índice de escolaridade.

4.3 Primeiras análises

A história de Belo Horizonte começa em 1701 com a fundação do Arraial Curral del Rei. Depois, quando tornou-se capital de Minas Gerais, passou a ser chamada Cidade de Minas, em 1897. Foi nomeada Belo Horizonte somente em 1901. A Figura 4.5 mostra a planta geral da até então Cidade de Minas, projetada pelo engenheiro Aarão Reis entre 1894 e 1897.

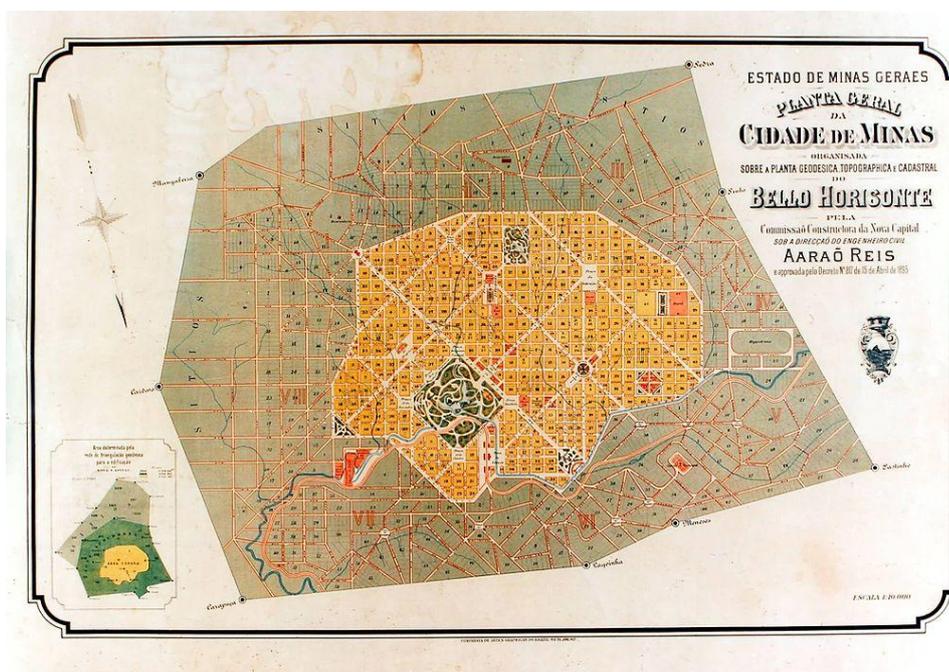


Figure 4.5: Planta Geral da Cidade de Minas

A seguir, está um trecho do relatório escrito por Aarão Reis, passando a ideia geral da organização da planta inicial de Belo Horizonte, aprovada pelo Decreto no 817 de 15 de abril de 1895: “Foi organizada, a planta geral da futura cidade dispondo-se na parte central, no local do atual arraial, a área urbana, de 8.815.382 m², dividida em quarteirões de 120m x 120m pelas ruas, largas e bem orientadas, que se cruzam em ângulos retos, e por algumas avenidas que as cortam em ângulos de 45 graus. Às ruas fiz dar a largura de 20 metros, necessária para a conveniente arborização, a livre circulação dos veículos, o tráfego dos carros e trabalhos da colocação e reparações das canalizações subterrâneas. Às avenidas fixe a largura de 35 metros, suficiente para dar-lhes a beleza e o conforto que deverão, de futuro, proporcionar à população (...)”.

A cidade foi dividida inicialmente em três principais zonas: a área central urbana, a área suburbana e a área rural. Podemos observar claramente esta divisão na planta mostrada na Figura 4.5. A Avenida do Contorno delimita a área central. Esta área receberia toda a estrutura urbana de transportes, educação, saneamento, assistência médica, abrigaria os edifícios públicos dos funcionários estaduais, além de possuir a maior parte dos estabelecimentos comerciais.

Esta seção tem como objetivo mostrar as primeiras análises feitas a partir de uma amostra de dados. Essa amostra consiste em uma parte da base de dados, composta por cerca de 25.000 árvores da região Centro-Sul de Belo Horizonte, dentro da região delimitada pela Avenida do Contorno. Esta região foi escolhida por ser uma das regiões mais antigas da cidade e, conseqüentemente, muitas árvores que se encontram nessa região são muito antigas e podem ser mais propensas à queda.

Foi obtida uma lista de 127 árvores perdidas no ano de 2017 dentro da região Centro-Sul. Essa lista de árvores foi gerada a partir de dados de outro setor da Secretaria Municipal de Meio Ambiente (SMMA) de Belo Horizonte, que é responsável pelo recolhimento de árvores perdidas ou caídas, não sendo um elemento do banco de dados de arborização nem o resultado de sua atualização, como seria desejável. A partir dos dados operacionais de recolhimento de árvores, foi necessário identificar as árvores suprimidas na base de dados do inventário utilizado no trabalho. Esse mapeamento foi feito através da geocodificação do endereço designado para o recolhimento na lista das árvores recolhidas e com o resultado deste processo, foi feita uma consulta espacial para descobrir a árvore do inventário mais próxima desse ponto. Todo esse processo reforça a dificuldade em encontrar uma base atualizada e unificada com as informações sobre as árvores da cidade, mesmo pouco tempo após o esforço de obtenção dos dados originais em campo.

O principal objetivo desta etapa é utilizar a mineração de dados para encontrar anomalias, padrões e correlações em grandes conjuntos de dados para prever resultados, gerando assim novos subconjuntos de dados que forneçam *insights* para solucionar problemas relacionados com a arborização urbana como, por exemplo, a queda de árvores.

4.3.1 Regras de Associação

Uma das principais ideias desta etapa é avaliar se existe alguma regra de associação entre árvores perdidas e outros atributos, a fim de identificar possíveis causas de quedas de árvores. Para tal, várias análises foram feitas, variando os valores de suporte mínimo e confiança.

A Figura 4.6 mostra uma comparação entre o suporte relativo da consequência

gerada pela regra e a confiança. Podemos observar que o suporte relativo assume apenas dois valores nas duas primeiras imagens, isto ocorre, pois dois atributos (PODADA e CCGE) possuem uma frequência muito alta na base de dados. Logo, com os valores de suporte e confiança estipulados para estes casos, apenas estes dois atributos aparecem como consequência das regras geradas. Isto nos permite observar que a confiança é bem sensível em relação à frequência do lado direito da regra, aqui chamado de consequência. Ou seja, dado uma regra $X \rightarrow Y$, uma frequência muito grande de Y pode fazer com que a regra possua uma confiança alta, mesmo se não houver uma associação forte entre X e Y .

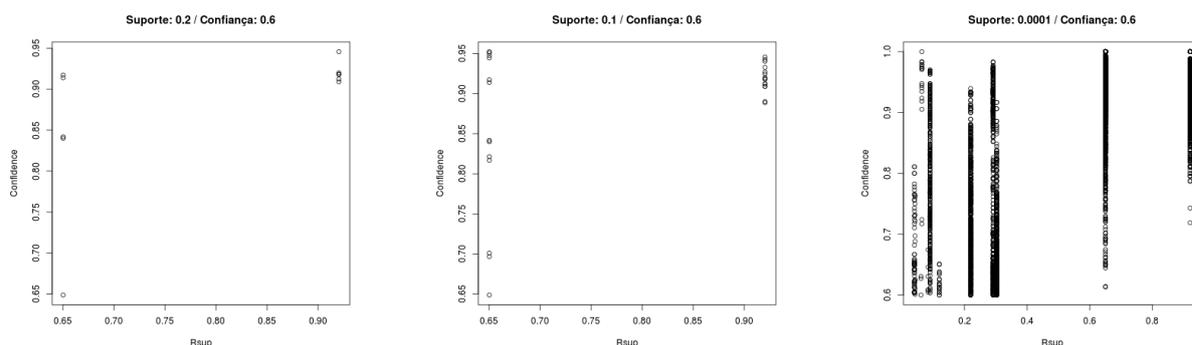


Figure 4.6: Relação entre suporte relativo e confiança

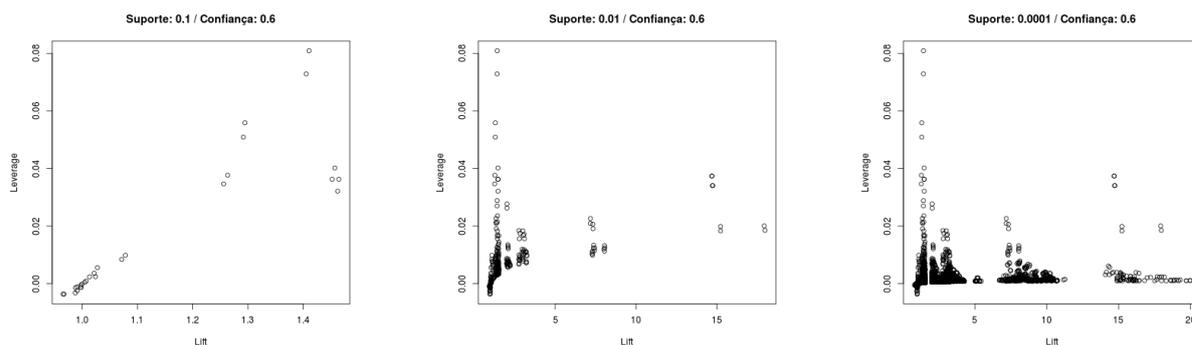
Com o intuito de extrair mais informações sobre as regras de associação e evitar os problemas de alta frequência do lado direito da regra, foram geradas novas métricas de interesse. As novas métricas analisadas foram o *Lift* e o *Leverage*.

O *Lift* indica o quão mais frequente é o lado direito da regra quando o lado esquerdo está presente, dado que, quanto maior seu valor, maior a relação entre os dois lados da regra. Caso o *lift* seja igual a 1, o lado direito ocorre com a mesma frequência independente do lado esquerdo ocorrer ou não, logo, os itens são independentes. Caso o *lift* seja menor que 1, o lado direito é mais frequente nas transações em que o lado esquerdo não ocorre.

O *Leverage* representa o número de transações adicionais cobertas pelos lados direito e esquerdo, caso os dois lados fossem independentes um do outro. Caso o *leverage* seja igual ou menor que 0, os dois lados da regra ocorrem juntos, podendo indicar que os dois lados são independentes. Deste modo, quanto maior o *leverage*, mais interessante será a regra.

A Figura 4.7 mostra a comparação entre o *lift* e o *leverage* das regras geradas variando o valor do suporte e confiança. Logo, observa-se que as regras mais interessantes são as regras que possuem altos valores de *lift* e *leverage*, ou seja, regras que se encontram próximo ao canto superior direito das Figuras mostradas em 4.7, justificando o uso das métricas *lift* e *leverage* para avaliar as regras descobertas, nos casos em que os valores de suporte são mínimos.

A Tabela 4.6 mostra algumas das regras mais interessantes que foram geradas.

Figure 4.7: Relação entre *lift* e *leverage*

Algumas regras já eram esperadas, como: BCPI, CCGS -> TCI; visto que, se a base da árvore possui presença de insetos, existe uma grande chance do tronco da árvore também possuir insetos. Outras regras, trazem descobertas interessantes, como: AR, PGC -> CCGE; visto que, em geral, quando há superficialidade da raiz e presença de galhos com casca existe uma grande chance da árvore possuir uma copa com galhos epicórmicos, estes são galhos podem indicar que as podas foram mal feitas e a superficialidade da raiz ou a presença de galhos com casca podem ter relação com tal evento. Para melhor entendimento de outras regras de associação seria necessário encontrar especialistas em árvores para entender fenômenos que podem acontecer dado a existência de uma regra de associação.

Regra de Associação	Confidence	Lift	Leverage
PERFILHAMENTO, BCA, PODADA ->BCBE	~0.600	~9.756	~0.003
BCPI, CCGS ->TCI	~0.726	~8.048	~0.013
CCGO, TCI ->CCGS	~0.705	~3.209	~0.007
RCE, PGC ->AR	~0.934	~3.204	~0.011
AR, PGC ->CCGE	~0.947	~1.457	~0.040

Table 4.6: Regras de Associação

O valor do suporte foi reduzido ao valor mínimo permitido pela ferramenta (0,0001), para tentar encontrar regras de associação que possuem o atributo PERECIDA e tentar identificar uma correlação entre a queda de árvores e outros atributos. A redução do suporte foi feita devido ao fato de que a base possui apenas 127 árvores precidas, o que corresponde a menos de 1% do total de árvores. Este número é pequeno, pois a base possui apenas as árvores que caíram no ano de 2017, porém já existe um esforço para a obtenção de mais dados junto à Prefeitura de Belo Horizonte.

A principal regra descoberta com relação às árvores precidas foi a regra PERECIDA -> CCGE, com confiança de aproximadamente 0.74 e um lift de 1.5; isto pode ser uma informação valiosa, dado que, conforme explicado anteriormente, galhos epicórmicos

podem ser provenientes de podas mal feitas. Logo, é possível que árvores tenham sofrido quedas por causa de podas mal feitas.

Embora o suporte não tenha sido relevante, visto que o padrão que estava sendo procurado (árvores perecidas) possui uma frequência muito pequena na base de dados, foram geradas várias regras de associações variando o valor do suporte e confiança para verificar a existência de outras regras que fossem interessantes e fazer descobertas além do padrão procurado.

Foram encontradas algumas regras relevantes, como a relação entre a superficialidade da raiz, a presença de galhos com casca e copas com galhos epicórmicos.

Outro insight fornecido por esta análise é que árvores perecidas, em geral, possuem copas com galhos epicórmicos. Isto pode ser um indício de que árvores tenham sofrido quedas por causa de podas mal feitas, visto que galhos epicórmicos podem ser provenientes de podas mal feitas. Entretanto, é necessário obter uma maior quantidade de árvores perecidas para confirmar tal fenômeno.

4.3.2 Agrupamento

O algoritmo de agrupamento escolhido foi o *K-Means*, visto que os dados não estão rotulados. Mesmo sabendo os rótulos das árvores que caíram, o objetivo é encontrar um grupo de árvores que possuem atributos similares às árvores perecidas e assim descobrir quais árvores ainda não caíram, mas que possuem características que as assemelham às perecidas, e portanto merecem um cuidado especial para evitar a queda.

Inicialmente o parâmetro escolhido (número de grupos) foi 2, na tentativa de separar as árvores que possuem chance de queda das árvores saudáveis, ou seja, dois grupos apenas. Os centróides iniciais foram gerados utilizando uma versão paralela do *k-means*, o *k-means++*. O uso dessa abordagem garante uma aproximação da solução ótima para encontrar os centróides iniciais, conforme exibido em [6]. Uma tolerância relativa para o término da execução do algoritmo para a soma das distâncias intra-clusters foi de 0.0001.

A Tabela 4.7 mostra o resultado da clusterização para $k = 2$. Podemos observar que houve uma boa separação das árvores perecidas (88.19% no primeiro cluster, e o restante no segundo cluster). Isto sugere que o segundo cluster descreve melhor as árvores com chance de queda, e todas as árvores que estão nesse grupo merecem uma atenção especial.

Podemos observar que os atributos escolhidos para realizar a etapa de agrupamento conseguem dividir a base de dados em um grupo de árvores potencialmente saudáveis e outro grupo com as árvores potencialmente não saudáveis. Isso indicaria que as árvores ainda vivas e que se encontram no cluster 1 necessitam de maior atenção, e uma avaliação

Cluster	Quantidade de árvores (% em relação ao total de árvores)	Quantidade de árvores perecidas (% em relação ao total de árvores perecidas)
0	22.716 (88.36%)	15 (11.81%)
1	2.994 (11.64%)	112 (88.19%)

Table 4.7: $K = 2$

mais rigorosa e presencial deve ser feita.

Em (a) da Figura 4.8 pode-se visualizar todas as 25.710 árvores analisadas neste estudo de caso, esta amostra possui apenas árvores que se encontram dentro da região delimitada pela Avenida do Contorno. Os pontos destacados na Figura 4.8 representam, respectivamente: (1) Parque Municipal Américo Renné Giannetti; (2) 12^o Batalhão de Infantaria Leve de Montanha; e (3) Palácio da Liberdade. Nestas regiões não existem árvores cadastradas, visto que o inventário possui apenas árvores em ruas e praças, não contemplando parques públicos ou outros locais privados. Já em (b) da Figura 4.8 pode-se observar a ampliação de uma parte da imagem vista em (a), onde são mostradas as árvores encontradas na região da Praça da Liberdade, uma das principais praças da cidade de Belo Horizonte.

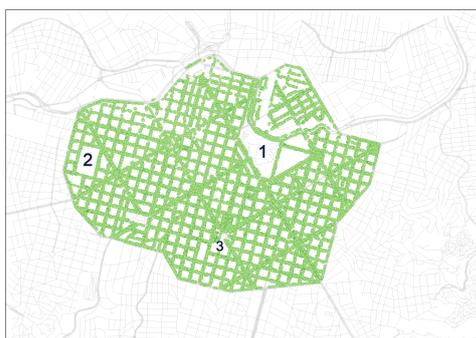


Figure 4.8: (a) Árvores dentro da Av. Contorno; (b) Árvores na região da Praça da Liberdade

A Figura 4.9 mostra as 2.994 árvores do cluster 1. Ao comparar esta imagem com a figura 4.8 podemos perceber a redução no espaço amostral de árvores a serem analisadas, tornando-se mais viável realizar uma triagem presencial para confirmar a validade das descobertas aqui descritas.

Por fim, a Figura 4.10 exibe um mapa de calor gerado com as árvores do cluster 1, ou seja, árvores que possuem uma chance maior de queda. Desta forma, é possível verificar regiões que necessitam maior atenção, pois possuem uma concentração maior de árvores com risco de queda.

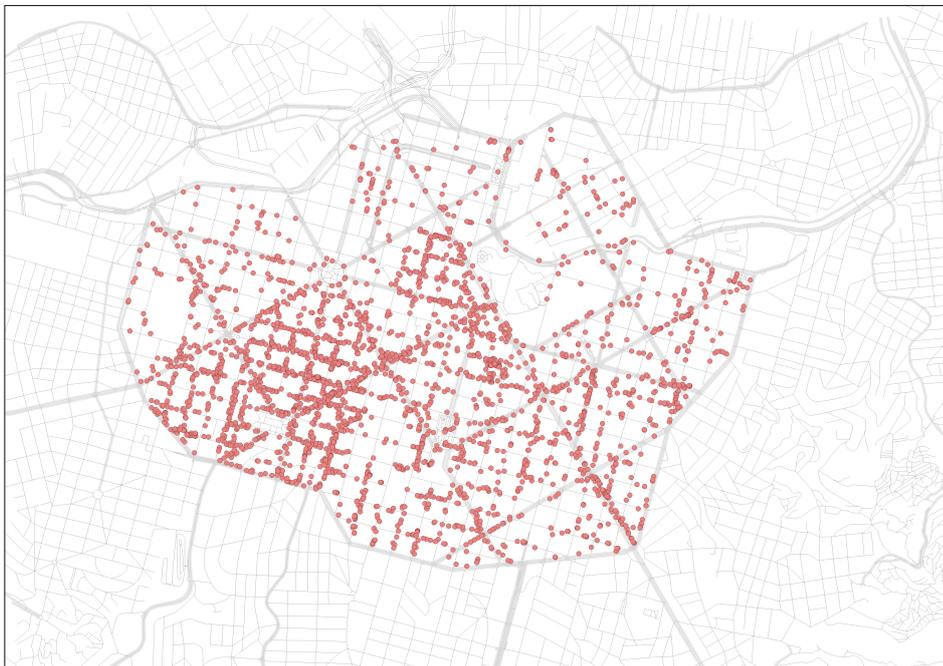


Figure 4.9: Árvores cluster 1



Figure 4.10: Mapa de calor das árvores com risco de queda

4.3.3 Considerações

A aplicação de técnicas de mineração de dados em um conjunto de dados existente sobre arborização urbana de uma grande cidade permitiu uma série de descobertas relacionadas com os indivíduos arbóreos analisados. As descobertas feitas ao aplicar a técnica de regras de associação mostram que existem atributos que possuem uma relação forte com árvores perecidas, como é o caso da presença de galhos epicórmicos em uma árvore. Enquanto as descobertas feitas ao aplicar a técnica de agrupamento sugerem que existe um conjunto de atributos capaz de separar árvores saudáveis das árvores com maior risco de queda, o que pode guiar o processo de inspeção e auxiliar em supressões preventivas visando evitar quedas inesperadas.

Chapter 5

VGI

Após as descobertas do capítulo anterior, notou-se a carência de uma base de dados unificada e padronizada para a reprodução em maiores escalas e confirmação dos resultados encontrados. Além disso, percebe-se que a base de dados atual está obsoleta, pois não é atualizada desde 2013, e várias modificações podem ter ocorrido nas árvores durante este tempo.

Para resolver este problema, foi criado um aplicativo que tem como objetivo manter os dados atualizados e concentrar as informações relativas às árvores de um município em uma única base. O processo de coleta é feito por usuários voluntários e que se identificam e se importam com o tema, o que define o conceito de *Volunteered Geographic Information* (VGI).

A Figura 5.1 descreve o fluxo de uso da aplicação e como é feito o recebimento e gerenciamento dos dados. As próximas seções trazem mais detalhes sobre o projeto, a implementação e o funcionamento da aplicação.

5.1 Projeto

O uso de técnicas de VGI não garante que os dados coletados sejam inteiramente confiáveis e também não garante uma cobertura completa dos dados esperados. Por isso, é indicado usar as técnicas de VGI quando os dados estão presentes no cotidiano dos cidadãos comuns, quando os dados estão bem distribuídos e em casos em que as coletas não exijam instrumentos específicos para serem feitas.

Os problemas de credibilidade de dados coletados através de técnicas de VGI são levantados por [19] e reforçam a necessidade de entender os motivos que levam as pessoas a contribuir com tais ferramentas. Dentre os principais pontos citados pelos autores estão: a dificuldade de se documentar as origens das fontes de informação, a inexistência de um indivíduo com conhecimento técnico específico no papel de responsável pela curadoria das informações e a falta de padrões mais rígidos de qualidade.

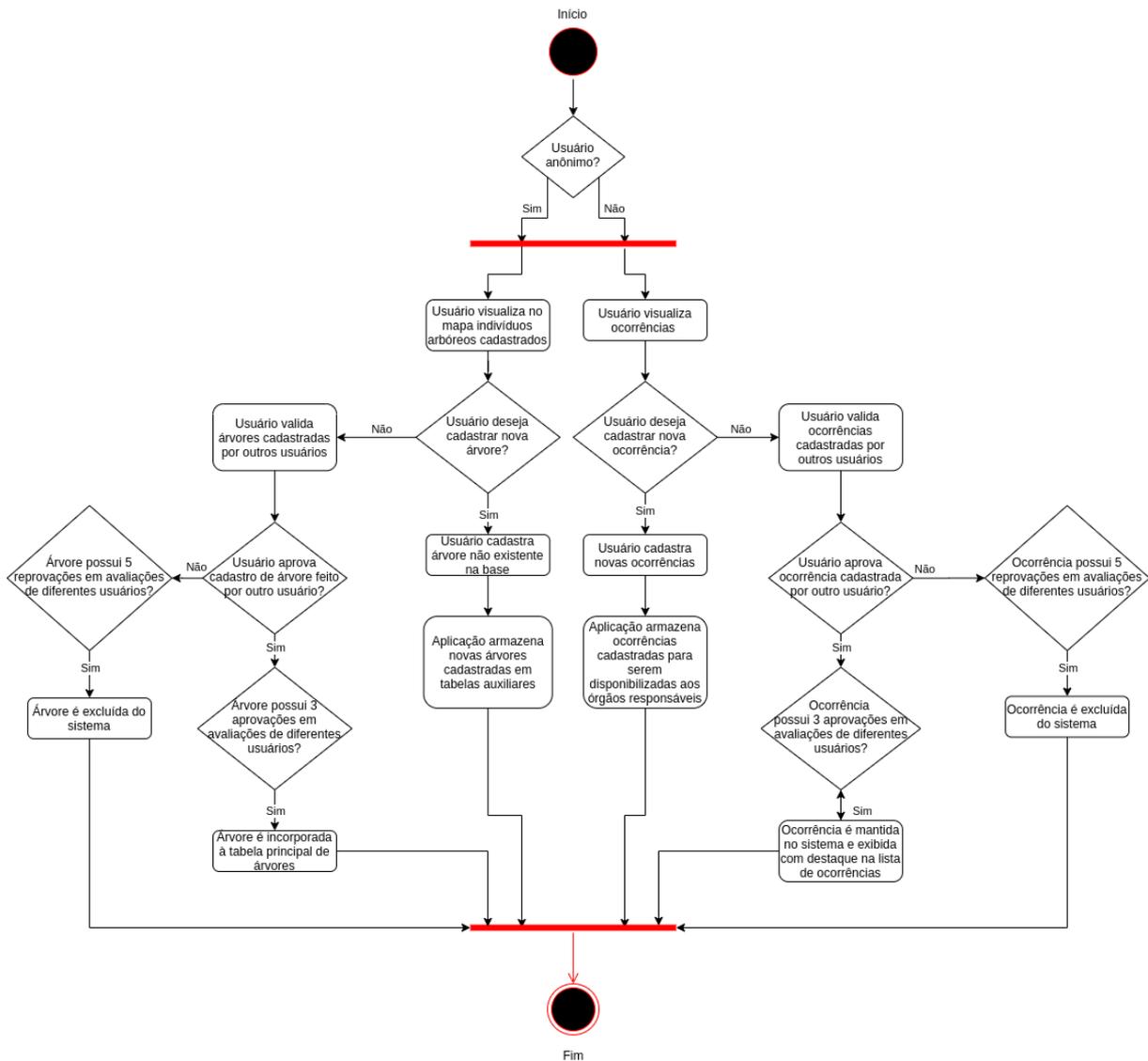


Figure 5.1: Fluxograma da aplicação

Existem diversas formas de contornar os problemas relacionados às técnicas de VGI citados anteriormente. A *gameificação* é uma estratégia utilizada com o intuito de incentivar a colaboração dos usuários através de desafios e recompensas [4]. Na temática de coleta de dados sobre arborização urbana a estratégia de *gameificação* pode não ser adequada, visto que a tendência é de que o cidadão se concentre em árvores situadas no entorno de sua casa, ou lugares que conhece. Logo, uma parcela significativa dos usuários pode possuir um espaço limitado de contribuição, e conseqüentemente, sua "pontuação" no processo de *gameificação* estará limitada à quantidade de contribuições.

Uma série de outros métodos sugere possibilidades para a identificação ou não dos usuários voluntários e estratégias para a validação das contribuições através de especialistas ou por outros usuários [23, 29]. Seguindo a abordagem de [23] podemos utilizar o próprio ambiente de *crowdsourcing* para validar as contribuições. Desta forma, os usuários

da ferramenta também fazem a manutenção e curadoria dos dados, além da coleta propriamente dita. Esta abordagem é chamada de verificação por pares e é adotada por várias ferramentas existentes, como *OpenStreetMap*¹, *WikiMapia*² e *ClickOnMap* (uma plataforma para o desenvolvimento de sistemas VGI, apresentada em [13]).

Observando os problemas e métodos apontados como estratégia para validação das contribuições em sistemas VGI e avaliação dessas estratégias aplicadas ao cenário da coleta voluntária de informações sobre arborização urbana, a ferramenta aqui apresentada utiliza a abordagem de verificação por pares, pois esta abordagem permite uma validação de custo menor, ao contrário das estratégias que utilizam especialistas para esta tarefa; além de não necessitar que os contribuintes naveguem por toda a malha existente, permitindo contribuições locais e dos lugares que estão presentes no seu cotidiano.

A ferramenta permite a parametrização das variáveis que irão definir a validação de uma contribuição. Desta forma, por padrão, a ferramenta necessita três validações por outros usuários para a confirmação do dado, e 5 validações negativas para a exclusão do dado.

Para facilitar a coleta em campo, a ferramenta foi desenvolvida para ser utilizada em um aplicativo móvel, permitindo que o usuário realize contribuições ao andar na rua e identificar ocorrências relacionadas à indivíduos arbóreos existentes na aplicação, cadastrar novos indivíduos e validar as contribuições feitas por outros usuários. Além do aplicativo móvel, a aplicação permite trafegar os dados coletados pelo aplicativo no dispositivo móvel e salvá-los em um banco de dados para persistência.

[15] descrevem uma arquitetura dividida em três camadas: *aplicação, serviços e persistência*. Essa abordagem define a separação entre os componentes da aplicação e permite que futuras implementações e extensões sejam feitas de maneira mais fácil, além de permitir a atualização de qualquer um dos componentes envolvidos de maneira independente.

A Figura 5.2 mostra o diagrama de componentes da aplicação aqui proposta, seguindo a divisão em três camadas mencionada anteriormente. A camada de aplicação é composta por um aplicativo móvel desenvolvido com uma tecnologia baseada em *Hypertext Markup Language* HTML e *Cascading Style Sheets* CSS, que posteriormente pode ser utilizada para a criação de uma aplicação web utilizando o mesmo código. A camada de serviços possui uma *Application Programming Interface* (API) utilizando o padrão *Representational State Transfer* (REST), que tem como objetivo controlar o acesso aos dados da camada de persistência por requisições vindas da camada de aplicação. Por fim, a camada de persistência é composta por um *Sistema de Gerenciamento de Banco de Dados* SGBD, responsável por gerenciar, manipular, persistir e organizar os dados. Mais detalhes sobre cada camada serão vistos na seção 5.2, dedicada à explicação do processo

¹<https://www.openstreetmap.org/>

²<https://wikimapia.org/>

de implementação.

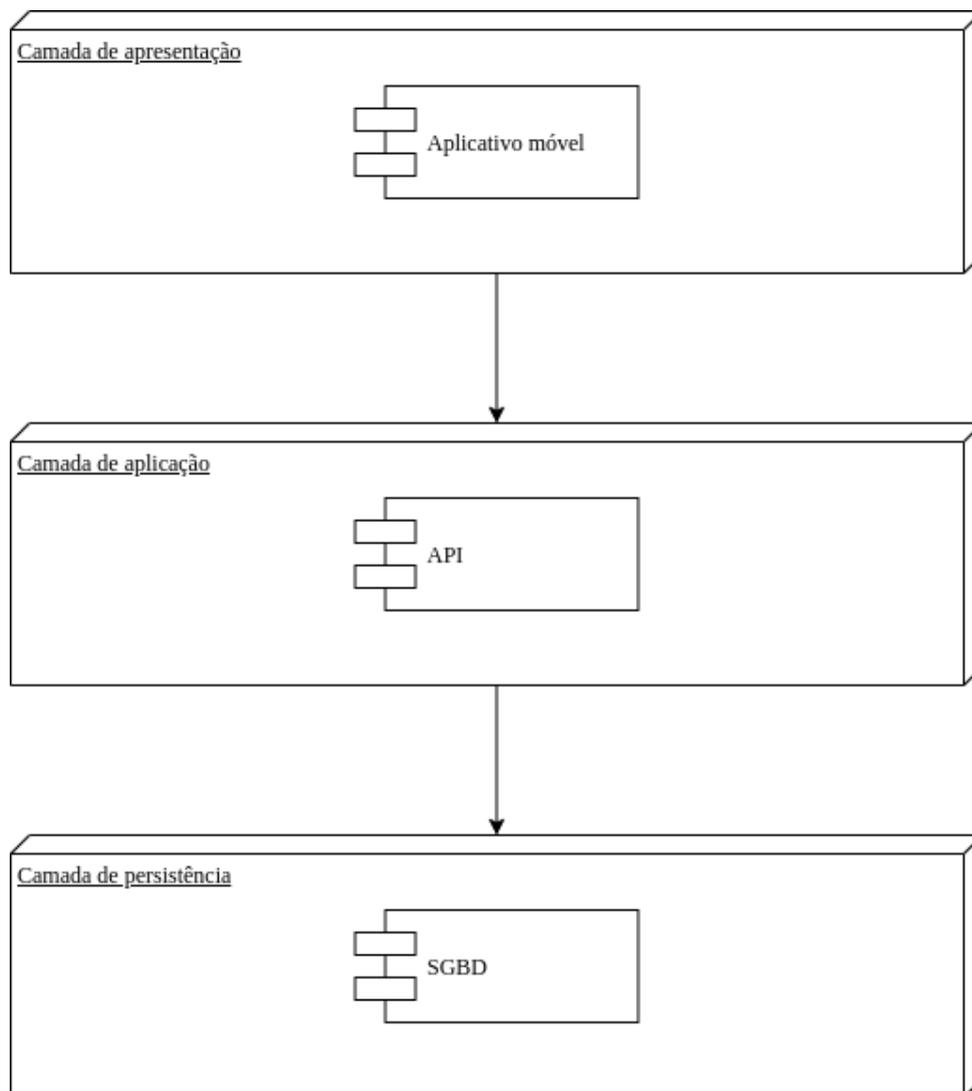


Figure 5.2: Diagrama de componentes

5.2 Implementação

Após as definições de projeto exibidas na seção 5.1, foi inicializado o processo de implementação da ferramenta aqui proposta. Inicialmente, foi criado um modelo de dados que contemple as entidades mapeadas a partir dos requisitos funcionais e requisitos não-funcionais da ferramenta. A Figura 5.3 exibe o modelo proposto com os atributos principais de cada classe.

O modelo *Object Modeling Technique for Geographic Applications* OMT-G foi es-

colhido para a modelagem dos dados. [10] definem o modelo OMT-G como um modelo de dados dotado de recursos para o projeto de bancos de dados e aplicações geográficas. O OMT-G parte das primitivas definidas para o diagrama de classes da Unified Modeling Language (UML), introduzindo primitivas geográficas com o objetivo de aumentar a capacidade de representação semântica daquele modelo e, portanto reduzindo a distância entre o modelo mental do espaço a ser modelado e o modelo de representação usual.

A utilização de um diagrama OMT-G permite que o mapeamento entre o projeto conceitual e a implementação física possa ser conduzido com mais segurança e com a preservação da semântica expressa no nível de abstração mais alto. Por isto, o diagrama visto em 5.3 foi elaborado para guiar e manter um padrão durante o desenvolvimento da ferramenta.

A classe *Usuário* é utilizada para persistir os usuários que possuem cadastro na ferramenta. Nesta classe também existe um usuário pré-cadastro no banco, que representa o usuário anônimo e é utilizado quando um usuário deseja acessar a ferramenta sem se identificar. A classe *Árvore* armazena as árvores "oficiais", que podem ser carregadas a partir de um inventário ou base existente e que será alimentada com as contribuições dos cidadãos após o processo de validação. Por sua vez, a classe *Minhas_árvores* possui registros sobre as árvores que determinado usuário "adotou", para que ele possa receber notificações e visualizar de forma mais rápida os registros de tais árvores.

A classe *Tipo_ocorrência* é uma classe de domínio dedicada à guardar os tipos de ocorrência que podem ser registradas pelos cidadãos. Esta classe foi inicialmente preenchida com os seguintes valores: *Queda de árvore*, *Árvore com necessidade de poda*, *Árvore em contato com a rede elétrica*, *Raízes superficiais* e *Árvore com presença de insetos*. A classe *Ocorrência* tem como objetivo armazenar as ocorrências relacionadas com determinada árvore, cadastradas pelos usuários da ferramenta. Enquanto a classe *Avaliação_ocorrência* visa registrar as validações de usuários sobre ocorrências cadastradas por outros cidadãos.

Por fim, as classes *Árvore_VGI* e *Avaliação_usuario* são responsáveis por armazenar as árvores novas cadastradas por contribuição voluntária e as validações por outros usuários sobre as novas árvores cadastradas, respectivamente. A Figura 5.3 apresenta um diagrama de classes OMT-G simplificado para o aplicativo.

O aplicativo foi desenvolvido utilizando o Quasar³, um *framework* baseado em Vue.Js que permite criar sites e aplicativos responsivos de forma híbrida, ou seja, o mesmo código é capaz de gerar um website ou um aplicativo móvel (Android ou iOS). Desta forma, além de possibilitar o desenvolvimento de um aplicativo móvel para ser executado nos principais sistemas operacionais de dispositivos móveis do mercado, o *framework* também possibilita que o código implementado seja aproveitado para a criação de uma aplicação web ou *desktop*, caso seja necessário.

³<https://quasar.dev/>

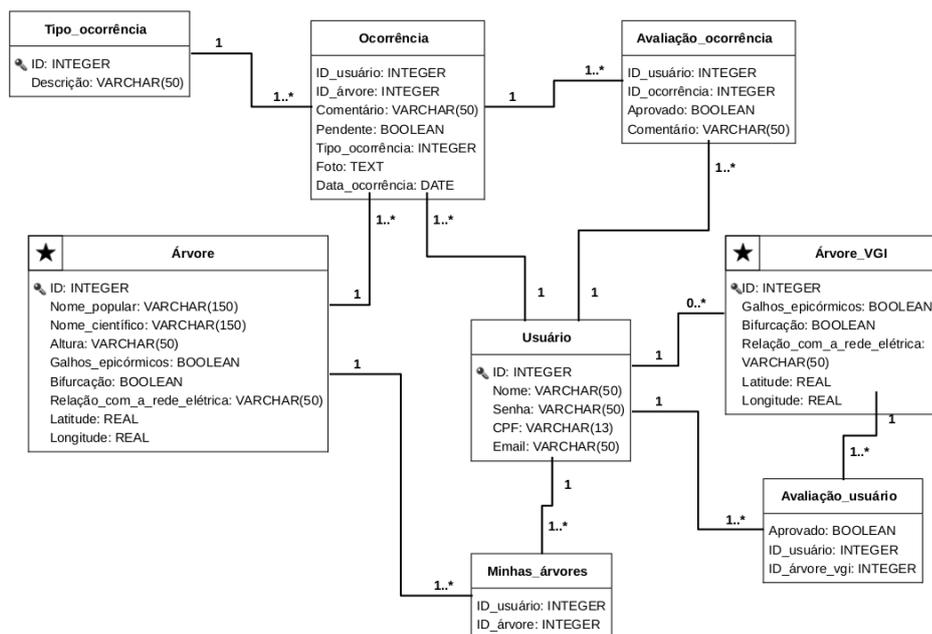


Figure 5.3: Diagrama de classes simplificado

A API foi desenvolvida utilizando Node.js ⁴, uma linguagem orientada a eventos com tempo de execução assíncrono e que foi projetada para a criação de aplicações escaláveis, conforme mostrado em [40]. Criar um ambiente Node.js e subir uma aplicação é uma tarefa que não exige muitos recursos computacionais em comparação com outras tecnologias mais tradicionais e também permite a criação de um ambiente escalável de forma simples e prática. Por isso, este *framework* foi escolhido para implementar a API aqui proposta.

A autenticação da aplicação é realizada através de um *JSON Web Token* (JWT), um método que define como transmitir e armazenar objetos JSON de forma compacta e segura entre diferentes aplicações. A API controla a criação, renovação e validação desse token de acordo com a necessidade. Toda requisição recebida pela API valida se o token existe no cabeçalho da requisição e se ele ainda é válido, ou seja, se o seu tempo de expiração ainda não foi atingido. Em caso de falha no processo de validação do token, é retornado o erro HTTP 401 *Unauthorized*, que indica que a solicitação não possui credenciais de autenticação válidas para essa requisição. Dessa forma, apenas usuários autenticados podem utilizar os recursos oferecidos pela API.

A Tabela 5.1 mostra uma lista com os métodos implementados na API com uma breve descrição. Mais detalhes podem ser analisados na documentação da API ⁵. A API e o banco de dados estão hospedados em um servidor da *Umbler* ⁶, que fornece a hospedagem de sites e aplicações e possui suporte ao Node.js e também ao MySQL, além

⁴<https://nodejs.org/en/docs/>

⁵<https://minhasarvores.com.br/apidoc/>

⁶<https://www.umbler.com/>

de possuir o deploy automático das aplicações e a criação de um ambiente escalável de maneira intuitiva.

Método	Caminho	Descrição
POST	/auth/register	Criação de um novo usuário
GET	/auth/user	Busca dados de um determinado usuário
POST	/auth/refresh	Renova o token de um usuário que está autenticado
POST	/auth/token	Cria um novo token após um usuário fornecer credenciais de acesso corretamente
POST	/ocorrencia/addOcorrencia	Insere uma nova ocorrência na base de dados
GET	/ocorrencia/ocorrencias	Lista todas as ocorrências existentes ordenadas de forma decrescente pela data de cadastro
GET	/ocorrencia/getTipo	Recupera o tipo de uma determinada ocorrência
GET	/ocorrencia/tipo	Recupera a lista de tipos possíveis de ocorrência
POST	/minhasArvores/add	Adiciona uma árvore à lista de árvores adotadas por determinado usuário
GET	/user/get	Busca informações de um determinado usuário
GET	/minhasArvores/get	Busca a lista de árvores adotadas por um determinado usuário
GET	/arvore/getArvore	Busca informações de uma determinada árvore
POST	/arvore/salvarArvore	Insere uma nova árvore
GET	/arvore/getArvores	Busca informações de um conjunto determinado de árvores
GET	/arvore/wfs	Busca o conjunto de árvores que se encontra dentro de um determinado polígono

Table 5.1: Métodos implementados na API

5.3 Funcionamento

Caso o usuário não deseje se identificar, é possível acessar o aplicativo de forma anônima e acessar quase todas as funcionalidades do sistema, com exceção do menu "Minhas Árvores". Logo, com o usuário anônimo, não é possível adotar árvores. A figura 5.4 mostra o processo de autenticação para acessar o sistema e o menu inicial que é exibido quando o acesso é efetuado com sucesso.

Com foco na usabilidade e com poucas ações possíveis, o aplicativo possibilita visualizar as árvores existentes na base de dados atual, bem como visualizar os atributos que podem ser interessantes para a população em geral. Os atributos escolhidos foram: nome popular, altura, nome científico, localização, base com insetos, copa com erva de passarinho, copa com galhos epicórmicos, relação de copa com cabos e o tipo da árvore (bifurcada ou não).

A Figura 5.5 mostra a visualização do mapa com as árvores já existente na base dados e a lista de atributos de uma árvore selecionada. Caso um usuário deseje visualizar os atributos de uma determinada árvore, basta clicar em cima dessa árvore no mapa que a janela de atributos com a opção de adotar uma árvore ou reportar uma ocorrência se abrirá.

Além disso, é possível registrar ocorrências para árvores existentes na base de dados ou até mesmo registrar ocorrência para árvores que ainda não se encontram na base. Estas ocorrências podem ser: relatar queda de um árvore, informar necessidade de

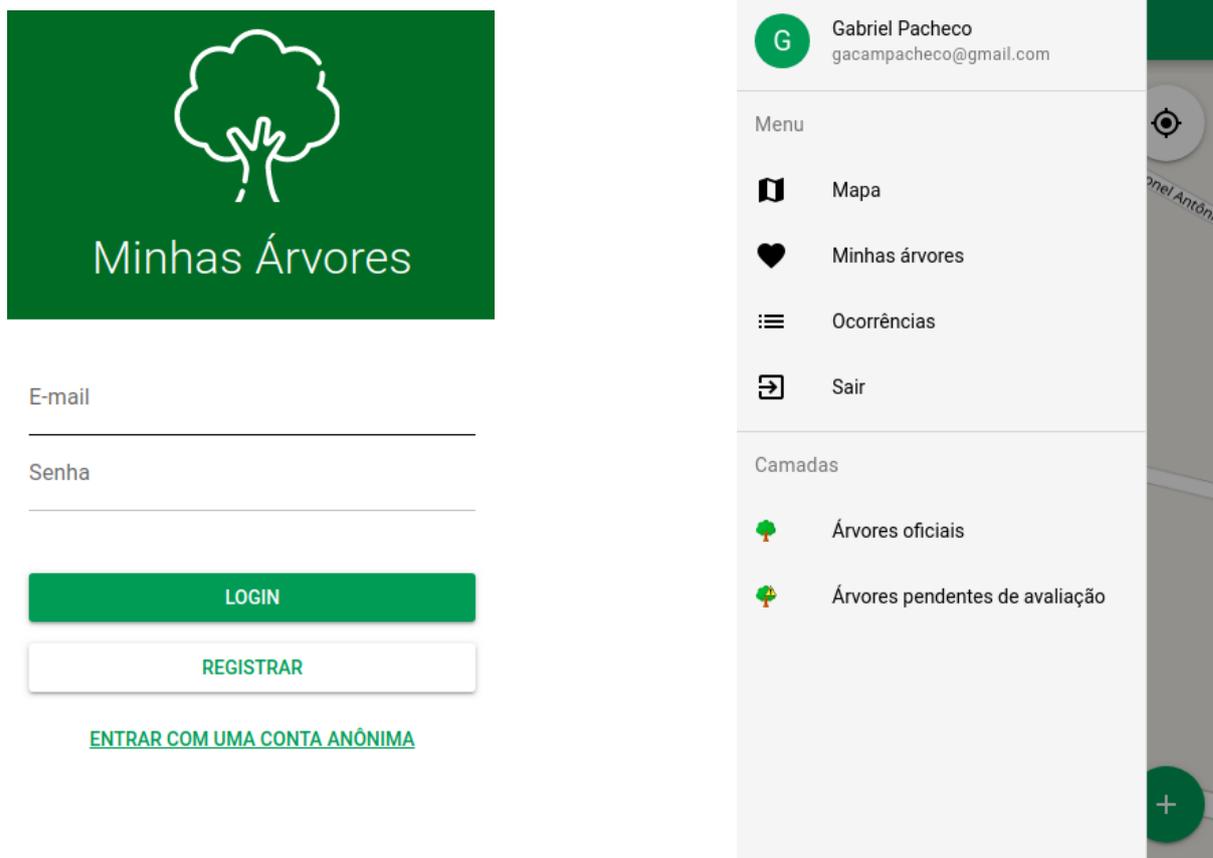


Figure 5.4: Autenticação e menu lateral

poda, informar contato com a rede elétrica, árvore com raízes superficiais ou informar a presença de insetos. Essas ocorrências são visualizadas em ordem decrescente, em relação a sua data de criação, através de um feed.

É possível “adotar” uma ou mais árvores, dessa forma o usuário pode receber notificações de ocorrências relativas às árvores adotadas por ele. Este recurso permite o usuário ter informações frequentes de árvores que estão, por exemplo, em sua rua, bairro ou próximo ao trabalho.

Também é possível cadastrar uma nova árvore, caso esta árvore não esteja na base atual. O intuito é que os usuários do aplicativo possam cadastrar árvores que não foram coletadas durante o projeto do SIA/BH, desta forma as árvores cadastradas através dessa funcionalidade poderão guiar o serviço público na coleta e atualização da base atual de dados.

Para acessar a funcionalidade de cadastro de árvore, basta clicar no botão “+” que se encontra no canto inferior direito no menu de Mapa. Ao clicar neste botão, será requisitado que o usuário clique no mapa e informe o local onde a árvore que está sendo cadastrada se encontra. Após este passo, serão feitas algumas perguntas sobre atributos da árvore em questão. Como os usuários podem ser leigos no assunto, os atributos foram mostrados em formas de desenhos ou croquis, para facilitar o entendimento.

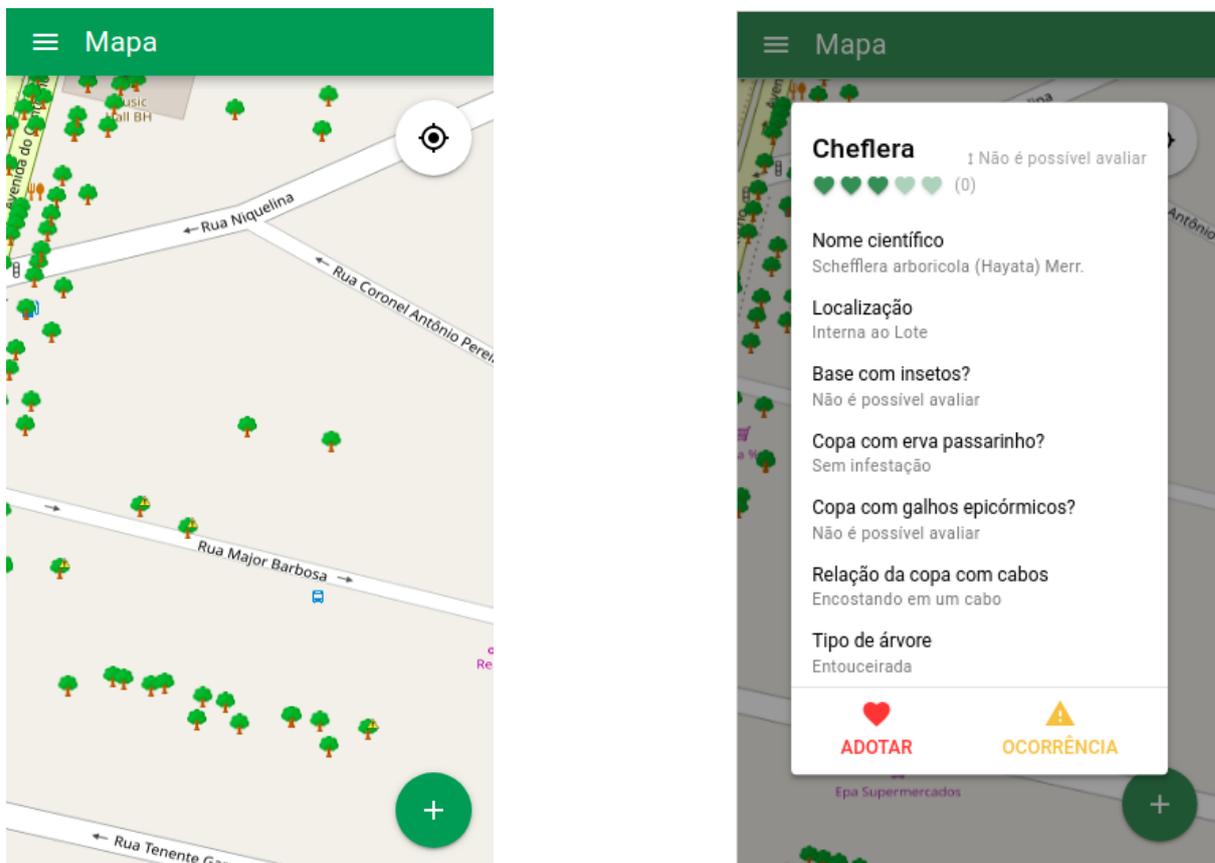


Figure 5.5: Mapa e atributos das árvores

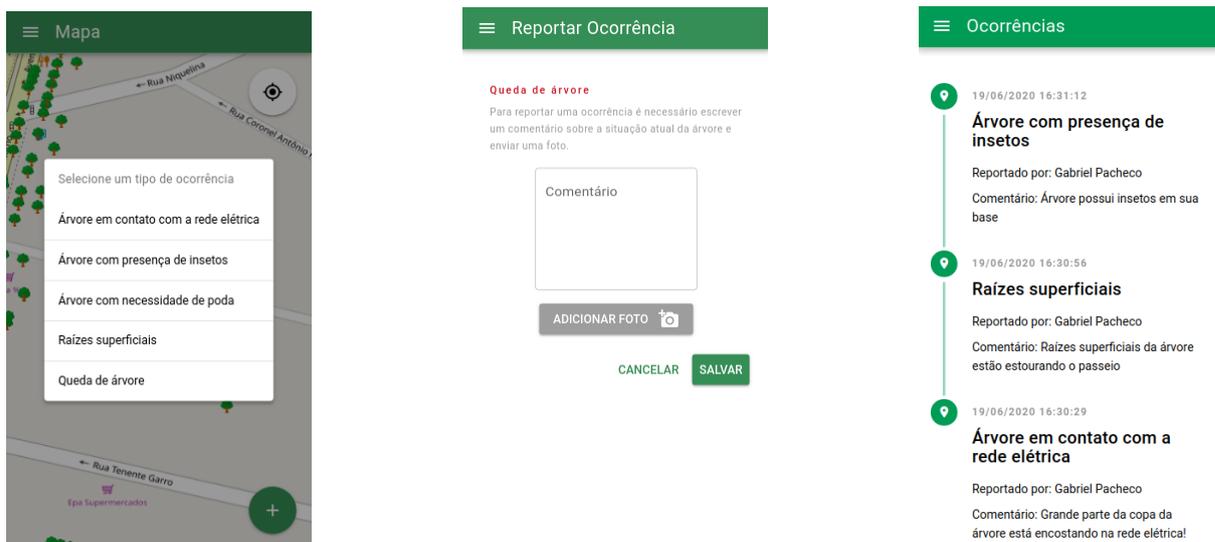


Figure 5.6: Cadastro e listagem de ocorrências



Figure 5.7: (a) identificação de bifurcação; (b) identificação de galhos epicórmicos

Na Figura 5.7 podemos observar que é requisitado informações sobre o estado de bifurcação da árvore, se ela é ou não bifurcada, e também se a árvore possui galhos epicórmicos ou não.

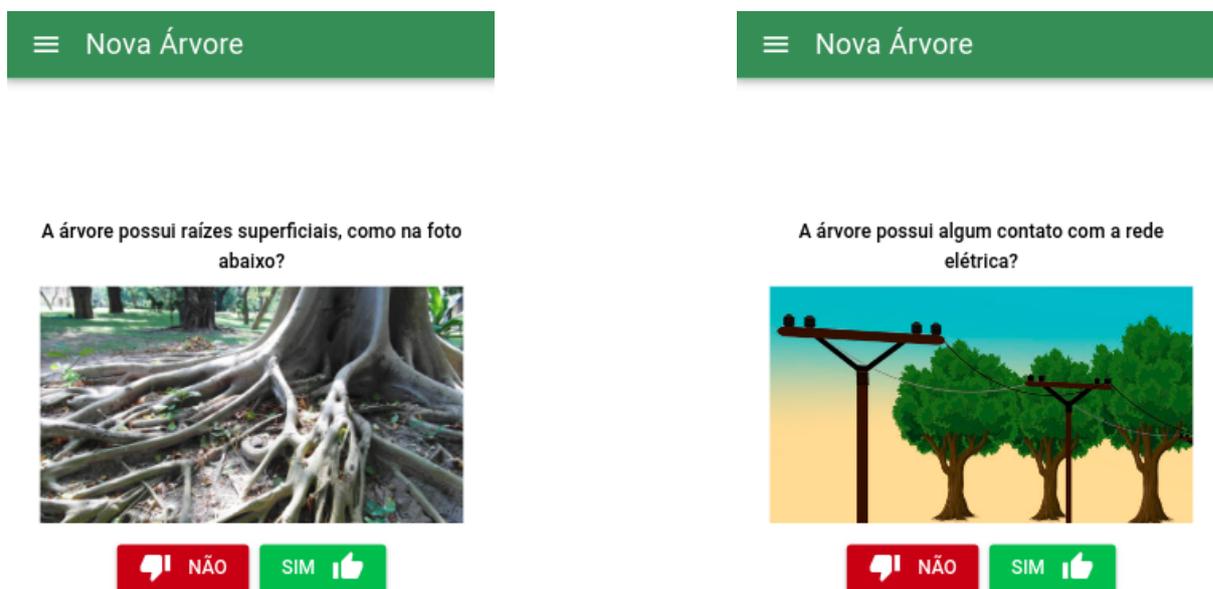


Figure 5.8: (a) identificação de raízes superficiais; (b) indicação de contato com a rede elétrica

As etapas do processo de cadastro de novas árvores exibidas na Figura 5.8 solicitam informações sobre a superficialidade das raízes e contato com a rede elétrica. Os atributos relacionados à bifurcação, presença de galhos epicórmicos e contato com a rede elétrica foram escolhidos por serem possíveis atributos que podem impactar diretamente no risco de queda de uma árvore, conforme mostrado na seção 4.3.2. Já o atributo sobre a presença de raízes superficiais foi solicitado, pois pode causar diversos estragos nos pavimentos das

calçadas e por isso, os órgãos públicos podem utilizar esta informação para buscarem as árvores que possuem esta característica e tomarem as devidas providências.

Os indivíduos arbóreos cadastrados por usuários através de contribuições voluntárias podem ser vistos no mapa com um ícone diferente das árvores oficiais (vide legenda das camadas exibida em (b) da figura 5.4). Desta forma, ao selecionar uma árvore que foi cadastrada por outro usuário é possível aprovar ou reprovar tal cadastro, conforme exibido em 5.9. Este processo de avaliação ou reprovação, caracteriza o processo de verificação por pares descrito anteriormente e que será utilizado para definir se uma árvore cadastrada voluntariamente por um usuário possui as informações corretas e se esta deve compor a tabela de árvores oficiais ou não. Logo, se uma árvore possuir três verificações positivas por diferentes usuários, esta será incorporada à base oficial de árvores. Caso algum registro possua cinco verificações negativas por diferentes usuários, o dado será excluído.



Figure 5.9: Verificação por pares

Desta forma, o aplicativo proposto possibilita que a população visualize a dis-

tribuição e informação dos indivíduos arbóreos próximos ao seu local de convívio. Além disso, é possível incluir novas árvores, relatar ocorrências relevantes em árvores existentes e visualizar ocorrência registradas de qualquer árvore na base de dados, de maneira simples e intuitiva. Também é possível avaliar indivíduos arbóreos cadastrados por outros usuários, a fim de realizar uma verificação por pares para controlar a qualidade dos dados coletados de forma voluntária.

Uma primeira versão da ferramenta foi implementada e se encontra pronta para ser testada. Os testes não foram possíveis em virtude da pandemia do novo coronavírus.

Além de ser utilizada pelos cidadãos, a ferramenta também pode ser utilizada por órgãos responsáveis pelo manejo da arborização urbana podendo ser utilizada para o registro de ocorrências relevantes como, por exemplo, podas e remoções de árvores.

Chapter 6

Conclusões

Este trabalho mostrou como utilizar mineração de dados para encontrar novos subconjuntos de dados que forneçam insights para solucionar problemas relacionados com a arborização urbana, como por exemplo, a queda de árvores e como a descoberta do conhecimento em dados existentes sobre árvores no espaço urbano de um grande centro auxiliaram na construção de uma ferramenta para a coleta voluntária de informações geográficas relacionadas com a arborização urbana.

Uma análise dos dados sobre arborização urbana de uma base existente da cidade de Belo Horizonte foi feita e mostrou características sobre a distribuição de indivíduos arbóreos no perímetro urbano da cidade. Também foi feita uma análise utilizando indicadores socioeconômicos da cidade em questão. Os resultados sugerem que regiões com melhores índices de renda média per capita, desenvolvimento humano e escolaridade possuem uma concentração maior de arborização urbana do que regiões menos desenvolvidas.

Os resultados descobertos com a aplicação de técnicas de mineração de dados sugerem que existem atributos mais frequentes em árvores que caíram e que podem indicar se uma árvore possui uma maior chance de queda ou não. Também foi visto que existem atributos capazes de separar as árvores que caíram das árvores saudáveis, e que as árvores que estão no mesmo grupo das que caíram, mas que ainda não sofreram queda, podem necessitar de uma atenção especial. Tais observações podem ser reforçadas utilizando dados de queda em períodos temporais posteriores, em um trabalho futuro decorrente das análises apresentadas aqui. Caso se verifique que os registros de queda posteriores ao fim de 2017 contêm uma predominância das árvores com atributos semelhantes aos das perdidas em 2017, será possível projetar um algoritmo preditor de quedas, o que ajudará a prefeitura em futuras operações de prevenção e de substituição de árvores.

Os resultados mostram que foram encontrados novos subconjuntos de dados que podem ser utilizados para a coleta voluntária de informações geográficas relacionadas com a arborização urbana. A contribuição dos cidadãos pode ser decisiva para a atualização dos dados de arborização urbana, indicando árvores suprimidas ou substituídas com base no conhecimento local das pessoas, além de alimentar a base de dados com alguns atributos que não necessitem de especialistas para o seu cadastro.

Foi apresentada uma ferramenta VGI, criada a partir dos resultados descobertos no

decorrer do trabalho, com o intuito de coletar informações sobre os principais atributos que impactam na saúde das árvores e tornar o processo de coleta voluntária e colaborativa de informações menos moroso e mais interessante para os usuários. Dados fornecidos voluntariamente podem enriquecer as indicações de queda e supressão utilizadas neste trabalho, que foram obtidas a partir de indicações indiretas, os registros de recolhimento de árvores perdidas.

Desta forma, a ferramenta VGI tem como objetivo realizar o cadastramento de novas árvores para possibilitar uma análise contínua da arborização urbana. Cidades que já possuem uma base de dados sobre a arborização urbana poderão utilizar a ferramenta para atualizar os registros existentes nessa base ou cadastrar novos registros que ainda não foram mapeados, como é o caso da cidade de Belo Horizonte. Com a base em constante atualização, a ferramenta ajudará na definição de diretrizes para pesquisa, capacitação e cooperação para o aprimoramento do espaço arbóreo urbano. A contribuição dos cidadãos também poderá auxiliar na definição de diretrizes para o manejo da arborização urbana, assim como de políticas e programas de administração e manejo. Com tais informações, os órgãos públicos poderão definir programas de produção e plantio de mudas mais adequadas às características dos logradouros públicos da cidade, e proporcionar um programa de educação ambiental para a população.

Portanto, este trabalho apresentou um processo de descoberta dos dados utilizando uma base existente sobre a arborização urbana de um grande centro e uma ferramenta de contribuição voluntária de informações para atualizar a base de dados existentes ou o cadastro inicial de inventário para cidades que ainda não possuam iniciativas nesse sentido. Com a atualização contínua dos dados de arborização urbana será possível refazer análises de mineração de dados para confirmar as descobertas feitas anteriormente, descobrir novos padrões, permitir novas descobertas, e conseqüentemente, sugerir novas diretrizes para a administração e manejo das árvores no espaço urbano.

Trabalhos futuros incluem a expansão da mineração de dados, análises espaciais de dados de árvores e disseminação da ferramenta VGI para envolver os cidadãos na coleta de dados. Analisar as árvores com risco de queda encontradas após aplicar a técnica de agrupamento, ampliando as análises das características de tais árvores e fazendo uma inspeção visual de uma amostra das árvores neste grupo para confirmar os resultados obtidos. A definição do processo de atualização dos dados oficiais com as informações coletadas através de cidadãos deve ser avaliada à medida em que a ferramenta passe a ser utilizada em larga escala. Utilizar as contribuições dos cidadãos para auxiliar na definição ou aprimoramento de diretrizes para o gerenciamento de árvores urbanas e áreas verdes, o desenvolvimento de uma interface para que os órgãos responsáveis possam visualizar as informações coletadas pela ferramenta VGI pode auxiliar nesse processo. Iniciativas educacionais, compreendendo informações sobre a adequação das espécies arbóreas às características dos espaços públicos, podem ocorrer simultaneamente. O trabalho pioneiro

da administração municipal de Belo Horizonte na criação e atualização de um inventário de árvores é aplicável a outras cidades, portanto as técnicas de análise e as iniciativas de VGI propostas neste projeto podem ser replicadas e reutilizadas.

Bibliography

- [1] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. Mining association rules between sets of items in large databases. In *Acm sigmod record*, volume 22, pages 207–216. ACM, 1993.
- [2] Rakesh Agrawal, Ramakrishnan Srikant, et al. Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB*, volume 1215, pages 487–499, 1994.
- [3] Ana Luísa Brito dos Santos Almeida et al. O valor das árvores: árvores e floresta urbana de lisboa. *Tese de Doutorado*, 1(1), 2006.
- [4] Vyron Antoniou and Christoph Schlieder. Participation patterns, vgi and gamification. *Proceedings of AGILE 2014. Presented at the AGILE*, pages 3–6, 2014.
- [5] Atlas do Desenvolvimento Humano no Brasil. Metodologia relativa à construção das unidades de desenvolvimento humano (udh), 2013.
- [6] Bahman Bahmani, Benjamin Moseley, Andrea Vattani, Ravi Kumar, and Sergei Vassilvitskii. Scalable k-means++. *Proceedings of the VLDB Endowment*, 5(7):622–633, 2012.
- [7] Pavel Berkhin. A survey of clustering data mining techniques. In *Grouping multidimensional data*, pages 25–71. Springer, 2006.
- [8] D. Biondi. *Árvores de rua de Curitiba: cultivo e manejo*. FUPEF, 2005.
- [9] João Henrique Bonametti. Arborização urbana. *Terra*, 2003.
- [10] Karla AV Borges, Clodoveu A Davis, and Alberto HF Laender. Omt-g: an object-oriented data model for geographic applications. *GeoInformatica*, 5(3):221–260, 2001.
- [11] JA Carvalho, João Carlos Nucci, and Simone Valaski. Inventário das árvores presentes na arborização de calçadas da porção central do bairro santa felicidade-curitiba/pr. *Revista da Sociedade Brasileira de Arborização Urbana, Piracicaba*, 5(1):126–143, 2010.
- [12] Fundação Biodiversitas Cemig. *Manual de Arborização*, 2011.

-
- [13] J. H. S. Câmara, L. F. M. Vegi, R. O. Pereira, Z. A. Geöcze, J. Lisboa-Filho, and W. D. de Souza. Clickonmap: A platform for development of volunteered geographic information systems. In *2017 12th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–6, 2017.
- [14] Ivan Dantas Coelho et al. Arborização urbana na cidade de campina grande-pb: Inventário e suas espécies. *Revista de Biologia e Ciências da Terra*, 4(2), 2004.
- [15] Clodoveu A Davis Jr, Hugo de Souza Vellozo, and Michele Brito Pinheiro. A framework for web and mobile volunteered geographic information applications. In *GeoInfo*, pages 147–157. Citeseer, 2013.
- [16] M. Demuzere, K. Orru, O. Heidrich, E. Olazabal, D. Geneletti, H. Orru, A.G. Bhave, N. Mittal, E. Feliu, and M. Faehnle. Mitigating and adapting to climate change: Multi-functional and multi-scale assessment of green urban infrastructure. *Journal of Environmental Management*, 146:107 – 115, 2014.
- [17] John F. Dwyer, E Mcpherson, Herbert Schroeder, and Rowan A. Rowntree. Assessing the benefits and costs of the urban forest. *J. Arbor.*, 18, 01 1992.
- [18] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. From data mining to knowledge discovery in databases. *AI magazine*, 17(3):37–37, 1996.
- [19] Andrew J. Flanagan and Miriam J. Metzger. The credibility of volunteered geographic information. *GeoJournal*, 72(3):137–148, Aug 2008.
- [20] William J Frawley, Gregory Piatetsky-Shapiro, and Christopher J Matheus. Knowledge discovery in databases: An overview. *AI magazine*, 13(3):57–57, 1992.
- [21] Asif Gill and Deborah Bunker. Crowd sourcing challenges assessment index for disaster management. In *AMCIS*, 2012.
- [22] Michael F. Goodchild. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4):211–221, Aug 2007.
- [23] Michael F Goodchild and Linna Li. Assuring the quality of volunteered geographic information. *Spatial statistics*, 1:110–120, 2012.
- [24] G.W. Grey and F.J. Deneke. *Urban forestry*. Wiley, 1986.
- [25] Usama M Fayyad Georges G Grinstein and Andreas Wierse. *Information visualization in data mining and knowledge discovery*. Morgan Kaufmann, 2002.
- [26] Flávio Eduardo Aoki Horita, Lívia Castro Degrossi, Luiz Fernando Gomes de Assis, Alexander Zipf, and João Porto de Albuquerque. The use of volunteered geographic

- information (vgi) and crowdsourcing in disaster management: a systematic literature review. In *AMCIS*, 2013.
- [27] IBGE. Características urbanísticas do entorno dos domicílios. *Censo demográfico*, 2010, 2010.
- [28] Brian A Johnson. High-resolution urban land-cover classification using a competitive multi-scale object-based approach. *Remote Sensing Letters*, 4(2):131–140, 2013.
- [29] Patrick Maué. Reputation as tool to ensure validity of vgi. In *Workshop on volunteered geographic information*. Citeseer, 2007.
- [30] Jennifer Mullaney, Terry Lucke, and Stephen J. Trueman. A review of benefits and challenges in growing street trees in paved urban environments. *Landscape and Urban Planning*, 134:157 – 166, 2015.
- [31] David J Nowak. Compensatory value of an urban forest: an application of the tree-value formula. *Journal of Arboriculture*. 19 (3): 173-177., 19(3), 1993.
- [32] Gabriel O. C. Pacheco and Clodoveu A. Davis Jr. Analyzing data on the tree coverage of a large city. In Jugurta Lisboa Filho and Antônio Miguel Vieira Monteiro, editors, *XX Brazilian Symposium on Geoinformatics - GeoInfo 2019, São José dos Campos, SP, Brazil, November 11-13, 2019*, pages 1–12. MCTIC/INPE, 2019.
- [33] Demóstenes Ferreira da. Pivetta, Kathia Fernandes Lopes; Silva Filho. Arborização urbana. *Boletim acadêmico*, 1:69, 09 2002.
- [34] Teresa Scassa. Legal issues with volunteered geographic information. *The Canadian Geographer / Le Géographe canadien*, 57(1):1–10, 2013.
- [35] Prefeitura Municipal de Belo Horizonte Secretaria Municipal de Meio Ambiente. Plano de mitigação de riscos advindos da arborização urbana da prefeitura de belo horizonte, 2018.
- [36] Prefeitura de São Paulo Secretaria Municipal do Verde e do Meio Ambiente. *Manual Técnico de Arborização Urbana*, 2nd edition, 2005.
- [37] Hansi Senaratne, Amin Mobasher, Ahmed Loai Ali, Cristina Capineri, and Mordechai (Muki) Haklay. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1):139–167, 2017.
- [38] James R Simpson and E Gregory McPherson. Potential of tree shade for reducing residential energy use in california. *Journal of Arboriculture*, 22:10–18, 1996.

-
- [39] Jan Tigges, Tobia Lakes, and Patrick Hostert. Urban vegetation classification: Benefits of multitemporal rapideye satellite data. *Remote Sensing of Environment*, 136:66–75, 2013.
- [40] Stefan Tilkov and Steve Vinoski. Node.js: Using javascript to build high-performance network programs. *IEEE Internet Computing*, 14(6):80–83, 2010.
- [41] Mohammed J. Zaki and Wagner Meira Jr. *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge University Press, May 2014.

Appendix A

Atributos na base de dados do SIA/BH

1. Grupo de atributos de identificação da árvore

- Coordenadas Geográficas
 - Na janela correspondente figuram as coordenadas do local no qual se marcou a representação da árvore
- Localização da árvore
 - Praça
 - Canteiro Central
 - Calçada
 - Área Remanescente
 - Faixa de rolamento
 - Interna ao lote
- Número do imóvel
 - Na janela vai figurar o número existente para o imóvel na frente do qual a árvore se localiza
- Nome do Logradouro
 - Na janela vai figurar a tipologia (rua, avenida, praça, etc.) e o nome oficial do logradouro conforme a base de dados da PRODABEL. Esse nome foi trazido da base georreferenciada
- Número de ordem
 - Na janela se fez o referenciamento à existência de mais de uma árvore lindeira a um lote. Quando existe apenas um lote em frente ao imóvel o número que figura é 1, que, no caso, se refere à primeira e única árvore aí representada. Para a correta identificação, utiliza-se como parâmetro para cada indivíduo arbóreo uma ordem que acompanha o sentido de crescimento da numeração na rua. No entanto, existem ruas nas quais a nu-

meraço não obedece sentidos de crescimento em determinados momentos ou podem usar mais de um sentido em trechos diferentes

- Complemento
 - Campo descritivo que serve para ajudar a localizar a árvore em frente a uma casa sem número, tomando-se cuidado com as observações feitas nos itens 2 e 3. Na janela faz-se a anotação de algo que se julga importante para ajudar a localizar a árvore, principalmente, quando não há número de identificação para o imóvel lindeiro a ela
 - Termos típicos que foram anotados nesse campo foram: “ao lado do número...”, “em frente a...”, “oposto a...”, “entre o número X e o número Y...”, seguidos de identificações de escolas, postos de gasolina, supermercados, etc.
- Tipo
 - Árvore
 - Toco
 - Muda
- Espécie
 - Nessa janela aparece o nome popular e o nome científico da espécie conforme uma listagem cadastrada

2. Grupo de atributos básicos

- Altura
 - Muda
 - Menor que 3 metros
 - De 3 a 6 metros
 - De 6 a 9 metros
 - Maior que 9 metros
 - Não é possível avaliar
- Rede Elétrica ou Ramal
 - Sem rede e sem ramal
 - Somente MT isolada
 - MT nua, BT isolada
 - Somente BT isolada
 - Alta Tensão
 - Somente BT com espaçador
 - Subterrânea

-
- Não é possível avaliar
 - Somente BT nua
 - MT nua, BT canaletada
 - MT nua, BT com espaçador
 - Ramal (sem rede)
 - MT nua, BT nua
 - Somente MT nua
 - MT protegida, BT canaletada
 - MT protegida, BT com espaçador
 - Somente BT canaletada
 - MT protegida, BT nua
 - Somente MT protegida
 - MT protegida, BT isolada
 - MT isolada, BT isolada
 - MT isolada, BT canaletada
 - MT isolada, BT nua
 - Posição da árvore em relação à rede ou ramal
 - Alinhado
 - Até 1 metro
 - De 1 a 2 metros
 - Acima de 2 metros
 - Em calçada oposta
 - Não é possível avaliar

3. Grupo de atributos sobre a copa da árvore

- Á árvore é
 - Não bifurcada
 - Bifurcada
 - Entouceirada
 - Não é possível avaliar
- Diâmetro Longitudinal (em metros)
 - Na janela está anotado um número que corresponde à medida do diâmetro observado na copa da árvore à época do recenseamento.
- Diâmetro Transversal (em metros)

-
- Na janela está anotado um número que corresponde à medida do diâmetro observado na copa da árvore à época do recenseamento.
 - Diâmetro de Galhos Invadindo a Faixa de Rolamento
 - Sem galhos até 5 metros de altura
 - De 1 até 10 centímetros
 - De 10 a 20 centímetros
 - Acima de 20 centímetros
 - Não é possível avaliar
 - Galhos interferindo a circulação em calçada
 - Sem galhos até 2,50 metros de altura
 - Galhos avançando até 1,00 metro sobre a calçada
 - Galhos avançando mais de 1,00 metro sobre a calçada
 - Não é possível avaliar
 - Copa com galhos ocos
 - Sim
 - Não
 - Não é possível avaliar
 - Copa com presença de fungos
 - Sim
 - Não
 - Não é possível avaliar
 - Copa com erva de passarinho
 - Sem infestação
 - Com infestação em até 20% da copa
 - Com infestação em mais de 20% da copa
 - Não é possível avaliar
 - Copa com folhagem rala ou de cor anormal
 - Não
 - Sim, em ponto restrito da copa
 - Sim, em toda a copa
 - Não é possível avaliar
 - Copa com poda unilateral
 - Não

-
- Sim, eliminando em menos de 40% em um lado
 - Sim, eliminando em mais de 40% em um lado
 - Não é possível avaliar
 - Copa com galhos secos
 - Sem galho importante seco
 - Com galho importante seco
 - Com predominância de galhos secos
 - Não é possível avaliar
 - Copa com lesão na casca
 - Sem lesão de casca
 - 1 a 30%
 - Maior que 50%
 - Não é possível avaliar
 - Copa com lesão na casca
 - Sem lesão de casca
 - 1 a 30%
 - Maior que 50%
 - Não é possível avaliar
 - Presença de galhos com casca inclusa
 - Sim
 - Não
 - Não é possível avaliar
 - Copa com galhos epicórmicos
 - Sim
 - Não
 - Não é possível avaliar

4. Grupo de atributos sobre o tronco da árvore

- Perfilhamento
 - Sim
 - Não
 - Não é possível avaliar
- Número de ramos perfilhados

- 1
- 2
- 3
- 4
- 5 ou mais
- Não é possível avaliar

- Diâmetro à altura do peito
 - Até 15 centímetros
 - De 15 a 30 centímetros
 - Acima de 45 centímetros
 - Não é possível avaliar

- Tronco com inclinação (Ressalta-se que por envolver uma dimensão angular difícil de ser verificada, a anotação vai do sentimento pessoal do recenseador)
 - Sem inclinação
 - De 1 a 20 graus
 - Acima de 20 graus
 - Não é possível avaliar

- Tronco inclinado sobre a faixa de rolamento
 - Sim
 - Não
 - Não é possível avaliar

- Tronco com lesão na casca
 - Anelamento
 - Até 5%
 - De 5 a 25%
 - Acima de 25%
 - Não é possível avaliar

- Tronco com grade
 - Sem grade
 - Com grade
 - Com grade inclusa
 - Não é possível avaliar

-
- Tronco com corpos estranhos
 - Não
 - Sim (sem interferência negativa relevante)
 - Sim (com interferência relevante)
 - Não é possível avaliar

 - Tronco com fungo
 - Sim
 - Não
 - Não é possível avaliar

 - Tronco com cavidade
 - Inexistente
 - De 1 a 60%
 - Acima de 60%
 - Não é possível avaliar

 - Tronco com indício de presença de inseto
 - Indício de cupim
 - Indício de outro inseto
 - Sem indício de inseto
 - Não é possível avaliar

5. Grupo de atributos da base do tronco

- Base com brotação epicórmica
 - Sim
 - Não
 - Não é possível avaliar

- Base com cavidade
 - Inexistente
 - De 1 a 60%
 - Acima de 60%
 - Não é possível avaliar

- Base com fungo
 - Sim

- Não
- Não é possível avaliar

- Elevação de solo e fissuras
 - Sim
 - Não
 - Não é possível avaliar

- Base com aterro
 - Sim, sem caixa de alvenaria
 - Sim, com caixa de alvenaria
 - Não, sem caixa de alvenaria
 - Não, com caixa de alvenaria
 - Não é possível avaliar

- Superficialidade do sistema radicular
 - Sem afloramento da raiz
 - Com afloramento da raiz e sem danos à calçada
 - Com afloramentos da raiz e com danos à calçada
 - Não é possível avaliar

- Base com indícios de presença de insetos
 - Indício de cupim
 - Indício de outro inseto
 - Sem indício de inseto
 - Não é possível avaliar

- Presença de raízes adventícias
 - Sim
 - Não
 - Não é possível avaliar

- Área livre ao redor da base
 - Justa
 - Extensão mínima menor que 15 centímetros
 - Extensão mínima maior que 15 centímetros
 - Não é possível avaliar

- Raízes cortadas ou estrangulantes
 - Sem raízes cortadas ou estrangulantes
 - Com raízes cortadas
 - Com raízes estrangulantes
 - Não é possível avaliar

- Lesão na casca
 - Até 5%
 - De 5 a 25%
 - Maior que 25%
 - Não é possível avaliar

- Tipo de piso
 - Solo nu
 - Vegetado
 - Pedras ou britas
 - Grelha ou arvoeiro
 - Cimentado
 - Outro tipo de pavimentação
 - Não é possível avaliar

6. Grupo de atributos da edificação lindeira à árvore

- Afastamento frontal da edificação
 - Menor que 1,50 metros
 - Maior que 1,50 metros
 - Edificação com marquise sobre a calçada
 - Edificação com balanço sobre a calçada
 - Não é possível avaliar

- Número de pavimentos da edificação
 - 1
 - 2
 - 3-5
 - Mais de 5
 - Não é possível avaliar

- Largura da calçada
 - Até 1,50 metros
 - De 1,50 metros a 2,20 metros
 - De 2,20 metros a 3,00 metros
 - Acima de 3,00 metros
 - Não é possível avaliar

- Calçada com indício de obra
 - Sim
 - Não
 - Não é possível avaliar

- Semáforo a menos de 2,00 metros de distância do tronco da árvore
 - Sim
 - Não
 - Não é possível avaliar

- Banca de revista, abrigo de ponto de ônibus ou equipamento fixo de porte similar a menos de 2,00m de distância do tronco da árvore
 - Sim
 - Não
 - Não é possível avaliar

- Árvore a menos de 5,00 metros de distância do poste mais próximo
 - Sim
 - Não
 - Não é possível avaliar