

Especificação do problema e seus aspectos éticos

Aline C. C. S. Azevedo
Instituto de Computação
UNICAMP
Campinas, Brasil
a189593@dac.unicamp.br

Beatriz C. Nascimento
Instituto de Computação
UNICAMP
Campinas, Brasil
b247403@dac.unicamp.br

Daniel G. Gratti
Instituto de Computação
UNICAMP
Campinas, Brasil
d214729@dac.unicamp.br

I. INTRODUÇÃO

No cenário atual, a avaliação e classificação de áreas com índices de criminalidade elevados têm emergido como um desafio crítico e complexo. A aplicação de técnicas de aprendizado de máquina permite auxiliar na análise desses índices através de uma vasta quantidade de dados. A capacidade de identificar e compreender essas regiões desempenha um papel fundamental para a formulação de estratégias de segurança pública. No entanto, é necessário também considerar fatores sociais para que não contribua ainda mais com viéses discriminatórios.

Apresentamos uma proposta que visa aprofundar a análise dos índices de criminalidade nas diferentes regiões de São Paulo, adotando uma abordagem ética e enfatizando a compreensão transparente do modelo. Nosso objetivo é identificar quais fatores socioeconômicos exercem maior influência na categorização do nível de segurança de uma área específica. Nesse sentido, nosso estudo se propõe a desenvolver um sistema para estimar o índice de criminalidade em regiões específicas e, concomitantemente, investigar o perfil socioeconômico das áreas mais impactadas pela violência urbana.

II. TRABALHOS RELACIONADOS

Fizemos uma revisão bibliográfica com estudos que utilizam aprendizado de máquina no contexto de criminalidade, com o foco em 4 artigos que discutimos a seguir.

O artigo *CrimAnalyzer: Understanding Crime Patterns in São Paulo* [1] tem objetivo de apresentar uma ferramenta de análise visual, *CrimAnalyzer*, que permite aos usuários estudar o comportamento dos crimes em regiões específicas da cidade de São Paulo - Brasil. O sistema ajuda a identificar áreas focais e os padrões de crimes associados, mostrando como esses padrões mudam ao longo do tempo. O *CrimAnalyzer* foi desenvolvido para auxiliar especialistas em criminologia e enfrenta três desafios principais: flexibilidade para explorar regiões locais e compreender os padrões de crimes, identificação de áreas focais espaciais de crimes importantes para investigação e compreensão da dinâmica dos padrões de crimes ao longo do tempo.

Outro artigo relacionado a estudo sobre análise de crimes na cidade de São Paulo é o *CriPAV: Street-Level Crime Patterns Analysis and Visualization* [2], onde é apresentada uma nova metodologia para lidar com análise de padrões de crime espaço-temporais em níveis detalhados de rua. A abordagem

proposta consiste em dois componentes principais: um mecanismo estocástico para identificar concentrações prováveis de crimes e um mecanismo de aprendizado profundo para mapear séries temporais de crimes em um espaço cartesiano, permitindo a identificação de locais com comportamento semelhante. Esses componentes são integrados na ferramenta *CriPAV*, uma plataforma *web* de análise e visualização de padrões de crimes. *CriPAV* foi desenvolvida em colaboração com especialistas da área e validada por meio de estudos de caso com dados reais de crimes em São Paulo. Os experimentos demonstram que o *CriPAV* é eficaz em identificar padrões, incluindo locais de crimes pouco intensos, mas altamente prováveis, e locais distantes com padrões de crimes similares.

O estudo *Exploring counterfactual antecedents to reduce criminality in Rio de Janeiro* [3] concentrou-se na análise do impacto de variáveis socioeconômicas e urbanas nas taxas de criminalidade na cidade do Rio de Janeiro. Utilizando dados por bairro, foram consideradas as taxas per capita de três categorias de crimes (contra pedestres, estabelecimentos comerciais e veículos), além de variáveis socioeconômicas e urbanas. A análise identificou áreas com maiores índices de cada tipo de crime e realizou cenários contrafactuais para locais específicos. Os resultados mostraram que tipos distintos de crimes ocorrem em diferentes partes da cidade. As análises contrafactuais revelaram variações significativas dependendo do bairro.

O estudo *Crime Analysis Through Machine Learning* [4] aborda a previsão de crimes usando aprendizado de máquina para análise de dados criminais de Vancouver ao longo de 15 anos com duas abordagens de processamento de dados distintas. Foram empregados modelos preditivos, incluindo o método dos *k*-vizinhos mais próximos e árvores de decisão impulsionadas. Os resultados demonstram que a previsão de crimes em Vancouver alcançou entre 39% e 44% de acurácia por meio desses modelos.

O artigo da ferramenta *CrimAnalyzer* [1] é o mais semelhante da proposta do nosso projeto, que também será focada na cidade de São Paulo. No entanto, trataremos de uma abordagem voltada para explicabilidade do modelo de aprendizado de máquina para possibilitar o entendimento e gerar possíveis soluções para problemáticas éticas que possam surgir.

III. DESCRIÇÃO DO PROBLEMA

Este projeto tem dois principais objetivos. O primeiro é modelar o problema de violência urbana a partir de relatos como boletins de ocorrência e questionários feitos a jovens em idade escolar sobre a percepção de violência em determinadas regiões de São Paulo. Em primeiro momento utilizaremos estes relatos de violência para rotular cada setor socioeconômico da capital paulista de acordo com o grau de violência, através da probabilidade dela ocorrer novamente num intervalo de tempo. O segundo é, utilizando recursos de explicabilidade, estudar o perfil social das áreas mais afetadas com violência urbana para responder à pergunta: “Quais são os fatores socioeconômicos que mais refletem as regiões mais afetadas pela violência urbana?”.

Esse projeto poderá ser utilizado para direcionar melhor os recursos financeiros e humanos objetivando reduzir os níveis de violência em regiões mais afetadas. Observando indicadores socioeconômicos detectados pelo modelo como determinantes no índice de criminalidade da região, é possível intervir com diferentes estratégias nestes espaços de modo a reduzi-la.

Os fatores socioeconômicos presentes em nosso dataset são importantes quando se tenta classificar uma área de acordo com seu nível de periculosidade, pois estão ligados a construções históricas que culminaram em nosso status quo. No entanto, nosso modelo pode incorrer em classificações erradas de áreas como perigosas ou não. Além disso, não sabemos se os participantes do questionário estavam cientes da utilização desses dados para tal classificação das regiões em que se encontram. A percepção de violência dos questionários pode ser enviesada por fatores sociais como racismo, machismo e econômicos. Isso pode resultar em uma distribuição injusta de recursos para esses lugares.

Escolhemos abordar este problema ético sob a ótica utilitarista. Esta é uma abordagem clássica para decisões com consequências que influenciam em grupos grandes de pessoas, pois nos instrui a determinar o quanto bem e mal serão produzidos por nossa ação [5]. O utilitarismo nos dá uma base clara para formular e testar políticas. Segundo esta teoria, uma decisão é boa se promove o maior bem estar geral em relação às alternativas. Ela disponibiliza uma abordagem flexível e orientada a resultados para a tomada de decisões morais. No caso do nosso problema, utilizaremos as métricas de desempenho do modelo e técnicas que promovam *fairness* para garantir que o menor número possível de pessoas sejam prejudicadas pelas limitações do algoritmo e pelos vieses contidos nos dados.

Nossa proposta busca desenvolver um sistema de aprendizado de máquina voltado para um problema de classificação, visando controlar vieses discriminatórios e mantendo uma boa performance. Para isso, levaremos em consideração possíveis problemas de segurança e privacidade, ao mesmo tempo que consideraremos cuidadosamente variáveis sensíveis, para atenuar qualquer tipo de prejuízo a grupos menos favorecidos. Essas variáveis serão tratadas de maneira apropriada para a problemática em questão.

Além disso, temos a intenção de explicar o funcionamento intrínseco do modelo, bem como das correlações que realiza para chegar em determinada saída. Com essa abordagem, acreditamos que nosso modelo será mais confiável, mitigando possíveis vieses sociais e, portanto, tornando-o apto para ser adotado por entidades públicas para que possa ser considerado em estratégias visando promover maior segurança em determinadas regiões. Com isso, incentivaremos também a colaboração da população no fornecimento contínuo de informações para a manutenção dos dados utilizados pelo sistema.

Tendo em vista esses aspectos, esperamos contribuir significativamente para a melhoria social, especialmente em áreas onde os índices de criminalidade apresentam-se mais elevados.

IV. PLANEJAMENTO

Para este trabalho utilizaremos duas bases de dados da cidade de São Paulo, uma de dados criminais providos pelo Núcleo de Estudos da Violência da USP, em São Paulo, que contém respostas de um questionário dado a estudantes de escolas da capital com perguntas envolvendo diversas faces de violência que os alunos vivenciassem, além de dados de boletins de ocorrência dos crimes de roubo e furto entre os anos de 2006 e 2017.

A segunda base de dados consiste de 21 dados socioeconômicos e de infraestrutura urbana, que caracterizam o perfil social de uma dada região principalmente obtidos através do censo brasileiro realizado pelo Instituto Brasileiro de Geografia. Nossa proposta é verificar se os dados socioeconômicos de uma região são suficientes para a previsão do padrão de crime observado.

Para isso, criaremos um modelo de aprendizado de máquina que, através dos dados de censo, seja capaz de prever o perfil de criminalidade local. Este perfil será extraído por outro modelo matemático, por meio da base dados de crime. Uma abordagem possível de ser feita é através da extração de características dos relatos e boletins de ocorrência, por exemplo, na Figura 1, pela determinação de parâmetros da distribuição de probabilidade do tempo entre ocorrências num setor, ou de forma automática, por outro modelo de aprendizado de máquina.

O objetivo então é aproximar a distribuição gerada pelos dados sociais a distribuição de dados criminais através de alguma métrica ou divergência apropriada, como uma regressão entre parâmetros ou vetores de características.

A elaboração desta proposta demanda a participação essencial de diversos profissionais capacitados. Recomenda-se a inclusão de um especialista em segurança pública, engenheiros e cientistas de dados qualificados e validadores de modelos para uma análise abrangente dos conceitos propostos. Além disso, a presença de engenheiros especializados em operações de aprendizado de máquina se faz necessária. Idealmente, a equipe deveria contar também com um perito em ética, como um sociólogo, a fim de garantir uma abordagem responsável. Adicionalmente, a perspectiva das comunidades locais pode

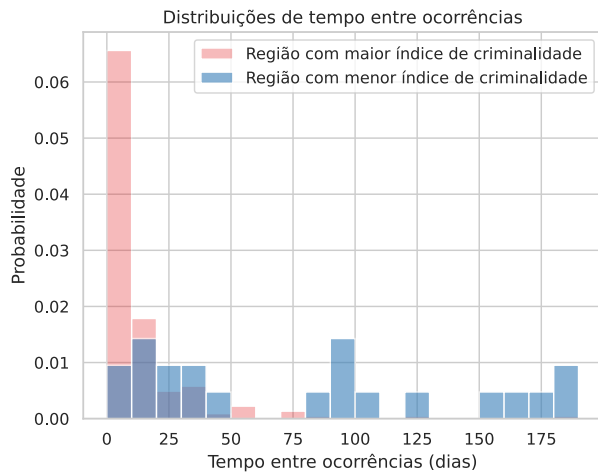


Figura 1. Diferentes níveis de criminalidade extraídos através dos dados temporais de boletins de ocorrência.

ser representada por um membro de uma associação de moradores das áreas envolvidas no estudo.

Com o objetivo de desenvolver um sistema direcionado à abordagem de uma questão sensível e complexa, como a criminalidade, o foco proposto é a explicabilidade do modelo, para identificar e atenuar problemas éticos, ampliando a transparência do sistema, a fim de estabelecer um nível mais elevado de confiança junto ao público-alvo.

V. LIMITAÇÕES E STATEMENTS

Essa tarefa não é isenta de desafios, uma vez que demanda a consideração cuidadosa de questões éticas, privacidade dos cidadãos e mitigação de discriminação, buscando equilibrar de maneira precisa a eficácia com a equidade. Nesse contexto, explorar abordagens que combinem rigor técnico com sensibilidade social torna-se decisivo para desenvolver soluções que promovam segurança e bem-estar de forma abrangente.

No desenvolvimento enfrentaremos o desafio de abordar os vieses presentes nas amostras coletadas, além da falta de dados existente. Uma vez que os dados de criminalidade funcionam como um indicador indireto do problema real, estamos sujeitos a certos vieses de localização, dado que áreas violentas podem existir onde o registro de boletins de ocorrência é escasso.

Dessa forma, buscaremos técnicas para reduzir os vieses que estão presentes no conjunto de dados e que podem surgir ao implementar o modelo. Para tal, procuraremos nos aprofundar em opiniões de especialistas de outras áreas do conhecimento que deveriam ser envolvidas no projeto. Essa é uma forma de termos uma perspectiva do problema uma vez que nosso grupo é formado por estudantes brancos da área de exatas, com ensino superior completo, que moram em uma cidade do interior de São Paulo em uma região favorecida.

Com esse propósito, nossa abordagem tem foco em explicabilidade, visando tornar os resultados acessíveis às pessoas que serão impactadas diretamente por eles e para que possamos assegurar uma contribuição para melhoria social.

REFERÊNCIAS

- [1] Germain Garcia, Jaqueline Silveira, Jorge Poco, Afonso Paiva, Marcelo Batista Nery, Claudio T Silva, Sergio Adorno, and Luis Gustavo Nonato. Crimanalyzer: Understanding crime patterns in sao paulo. *IEEE transactions on visualization and computer graphics*, 27(4):2313–2328, 2019.
- [2] Germain Garcia-Zanabria, Marcos M Raimundo, Jorge Poco, Marcelo Batista Nery, Cláudio T Silva, Sergio Adorno, and Luis Gustavo Nonato. Cripav: Street-level crime patterns analysis and visualization. *IEEE Transactions on Visualization and Computer Graphics*, 28(12):4000–4015, 2021.
- [3] Vitória Aquino Guardieiro. Exploring counterfactual antecedents to reduce criminality in rio de janeiro. Technical report, 2021.
- [4] Suhong Kim, Param Joshi, Parminder Singh Kalsi, and Pooya Taheri. Crime analysis through machine learning. In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pages 415–420. IEEE, 2018.
- [5] Alexander Mitov. Ethical use of artificial intelligence through the utilitarianism perspective. B.S. thesis, University of Twente, 2021.