

MO810 - Explicabilidade

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 a utilização de modelos de aprendizado de máquina não contribua ainda mais com vieses discriminatórios.

Apresentamos uma proposta que visa aprofundar a análise dos índices de criminalidade nas diferentes regiões de São Paulo, adotando uma abordagem ética. Nosso objetivo é treinar um modelo capaz de prever a criminalidade em regiões da cidade de São Paulo, utilizando como um *proxy* a quantidade de Boletins de Ocorrência (BO) feitos naquela região.

II. TRABALHOS RELACIONADOS

Realizamos uma revisão bibliográfica sobre estudos que empregam técnicas de explicabilidade em sistemas de aprendizado de máquina, predominantemente no contexto da criminalidade. O nosso foco recaiu sobre a análise detalhada de cinco artigos específicos, os quais discutimos a seguir.

O artigo "*Interpretable machine learning models for crime prediction*" aborda o uso de machine learning na previsão de crimes e destaca a importância da interpretabilidade desses modelos. Apesar de sua eficácia, os modelos de machine learning costumam ser considerados "caixas-pretas" devido à falta de transparência. Isso levanta desconfiança nos resultados. É ressaltada a necessidade de modelos transparentes e interpretáveis para prever crimes, considerando variáveis do ambiente. O estudo propõe métodos ante-hoc e post-hoc para melhorar a interpretabilidade dos modelos, dividindo a interpretabilidade em global e local. Essa interpretabilidade é importante para entender como o modelo toma decisões e aumentar a credibilidade dos resultados de previsão. Além disso, é proposto métodos para tornar esses modelos mais interpretáveis, destacando a importância de entender como eles tomam decisões, especialmente no contexto policial. O desafio é equilibrar a precisão e a interpretabilidade dos modelos. É apresentado o uso do modelo XGBoost com variáveis criminais e de ambiente, combinado com o método *Shapley additive*

explanation (SHAP) para interpretação, visando oferecer um modelo preciso e interpretável [1].

Já no artigo "*Mapping the risk terrain for crime using machine learning*" discute estratégias de policiamento focalizado em áreas de alta incidência criminal e a utilização de modelos de aprendizado de máquina para prever crimes futuros em locais específicos. Apresenta o modelo Random Forests (RF) como uma ferramenta para prever roubos em áreas específicas de Dallas, Texas, comparando-o com outras técnicas populares. Embora modelos de machine learning sejam eficazes na previsão de crimes, sua natureza opaca é frequentemente criticada. A pesquisa propõe o uso de estatísticas interpretáveis para avaliar esses modelos, destacando a importância de entender os fatores que influenciam as previsões criminais, como a distância de geradores de crimes, a influência da pobreza e a interação de fatores demográficos. Essas análises interpretáveis podem tornar esses modelos mais viáveis para pesquisadores interessados na relação entre crime e localidade [2].

Também encontramos o artigo "*Nonlinear relationships and interaction effects of an urban environment on crime incidence: Application of urban big data and an interpretable machine learning method*" que analisa a relação entre crimes urbanos e o ambiente urbano, enfatizando a importância de identificar áreas propícias a crimes para prevenção e políticas públicas. Teorias diversas sobre o tema são exploradas, mostrando que o ambiente urbano influencia a incidência de crimes de várias maneiras. Destaca-se a necessidade de compreender melhor as relações não-lineares e os efeitos de interação entre o ambiente urbano e o crime. Além disso, aborda-se o uso de técnicas de aprendizado de máquina para examinar essas relações, propondo o uso do SHAP para verificar esses efeitos. O estudo, dividido em cinco seções, explora dados urbanos para compreender a influência do ambiente urbano na incidência de crimes, fornecendo contribuições conceituais e metodológicas para campos como planejamento urbano, design urbano e criminologia ambiental, com implicações para políticas visando cidades seguras e sociedades sustentáveis [3].

O artigo "*Fairness by Explicability and Adversarial SHAP Learning*" aborda o crescente uso de algoritmos de machine learning em diversas indústrias e as preocupações relacionadas à interpretabilidade e justiça desses sistemas. Destaca-se a manifestação dessas preocupações na legislação da União Europeia, como na *General Data Protection Regulation*, enfatizando o direito individual à transparência e imparcialidade no

processamento de dados. O texto propõe o conceito de "Justiça pela Explicabilidade" como uma abordagem para avaliar a justiça de um algoritmo, relacionando-a à interpretabilidade do modelo e ao papel dos atores externos na auditoria do modelo. Para implementar essa abordagem, utiliza-se metodologias de interpretabilidade do modelo em conjunto com aprendizado adversarial. O estudo utiliza conjuntos de dados sintéticos e reais para ilustrar a viabilidade dessas abordagens em aplicações industriais, dividindo-se em seções que introduzem a notação, explicam os valores SHAP e discutem medidas de justiça estatística, culminando na apresentação dos resultados dos experimentos e nas conclusões [4].

III. METODOLOGIA

O objetivo desta análise é o treinamento de modelos explicáveis para o problema de análise de crimes segundo características socio-econômicas nas regiões de São Paulo e como estas características impactam no cenário criminal daquela área. Aqui, utilizaremos uma formulação e arquitetura de modelos diferente das utilizadas em trabalhos anteriores.

O desafio de prever crimes numa região poderia ser modelada como uma regressão para prever a distribuição dos boletins de ocorrência, por exemplo, no tempo. No entanto, regiões com baixo índice de criminalidade não possuem amostras de boletins de ocorrência capazes de se extrair uma distribuição que modele os BOs de forma satisfatória, ou até mesmo fiel. A forma que encontramos para minimizar este problema anteriormente foi de modelar o problema como um problema de classificação em regiões perigosas ou não, baseado apenas na quantidade de BOs registrados em dada região. No entanto, este tipo de redução carrega inúmeros vieses e não é justo para tomar qualquer tipo de decisão, principalmente devido a ambiguidade que um *hard threshold* causa ao diferenciar amostras cujo padrão é semelhante, mas o *ground-truth* induzido não é. Por exemplo, caso o *threshold* para determinar regiões perigosas seja ter mais de 1 BOs por dia, o modelo classificará regiões com média de 0.99 BOs por dia como seguras, apesar de potencialmente possuírem características parecidas.

Com isso, julgamos mais apropriado levar em conta a quantidade de boletins de ocorrência como um *proxy* direto do nível de criminalidade de uma região. Desta forma, formulamos o problema de entender padrões de crime através de um problema de regressão de contagem de boletins de ocorrência. Aproveitamos também os resultados do trabalho anterior para mitigar vieses providos da variável protegida de porcentagem de famílias de renda alta na região através da técnica de aumento de dados. Utilizamos uma aumento de 40% de nosso conjunto de dados com amostras sintéticas que mais se assemelham a amostras reais com valores da variável sensível reamostradas para descorrelacioná-la com o rótulo.

A. Modelo Interpretável

Curiel *et. al.* [5] utiliza uma modelagem de ocorrência de crime para medir a concentração de crime em áreas urbanas, de forma parecida com o problema aqui proposto.

A distribuição de crimes é modelada como uma distribuição Poisson-Binomial, em que cada indivíduo é selecionado por uma distribuição Binomial e, então, a distribuição condicionada ao indivíduo é modelada segundo uma distribuição de Poisson. Esta hipótese é atestada num conjunto de dados de vítimas de crimes em regiões urbanas do México, em que se conclui que a distribuição de crimes é extremamente desigual, com áreas e indivíduos estatisticamente imunes ao crime, e outras, em que a incidência criminal é mais forte.

Este fenômeno pode também ser observado no conjunto de dados criminais de São Paulo, onde há uma enorme população com baixos índices criminais, enquanto uma outra parcela sofre com maiores taxas, gerando dados extremamente desbalanceados, inflados em zero. Esta característica é um desafio para modelos estatísticos e pode gerar injustiças na explicação, para mitigar possíveis vieses no modelo realizamos uma análise exploratória entre modelos diretamente interpretáveis. Nossa análise focou em modelos mais clássicos, da classe dos modelos lineares generalizados, Regressão Linear, Regressão de Poisson e um modelo por Regressão Quantílica, que apesar de $\mathbb{E}[y | x]$ não ser linear, ainda atribui uma relação linear entre as features e o quantil que ela deve pertencer e, portanto, ainda pode ser interpretada diretamente.

Tendo ainda em vista a disparidade de zeros que o modelo contém, propomos ainda uma modificação dos modelos selecionados anteriormente ao acoplarmos o modelo a um classificador, que deve prever se uma amostra possui rótulo zero, caso o classificador determine o oposto, só então o regressor predirá qual o valor do rótulo. Essa classe de modelos é chamada de Modelos inflados em zero, e utilizamos modelos deste tipo compostos de um classificador linear e um modelo de regressão, conforme propostos anteriormente.

B. Modelo não-interpretável

Como um modelo complexo não diretamente interpretável escolhemos árvores de decisão treinadas com Gradient Boosting baseado em histograma. Este tipo de modelo é capaz de receber diferentes objetivos, incluindo os usados para regressão de Poisson e Quantílica, o que possibilita uma comparação direta com os modelos interpretáveis analisados.

Modelos baseados em *ensemble*, como árvores de decisão treinadas com Boosting, não podem ser diretamente interpretáveis devido a sua natureza de vários pequenos estimadores que aprendem apenas uma parcela da informação. Para que possamos prover uma explicação para a saída destes estimadores utilizaremos métodos de explicabilidade aproximada, em que o funcionamento exato do modelo não é explorado, mas apenas a relação que as saídas possuem com as entradas, inclusive para entradas que não estejam originalmente no conjunto de dados.

Para isso, utilizaremos dois métodos, um local e outro global. Para o método local utilizaremos o método de Explicações Aditivas de Shapley, SHAP (*Shapley Additive Explanations*). Esta técnica é uma abordagem para explicar as predições de modelos de aprendizado de máquina de forma transparente e interpretável. SHAP é baseado nos valores de Shapley

da teoria dos jogos cooperativos, que foram introduzidos por Lloyd Shapley e refletem as contribuições individuais a cada recurso (característica) em uma previsão específica. Ele considera todas as combinações possíveis de características e calcula a contribuição média de cada recurso para todas essas combinações. Isso é crucial para garantir que o mérito de cada recurso seja distribuído de maneira justa, levando em consideração as interações complexas entre eles.

O resultado final do SHAP é uma explicação transparente para cada previsão, indicando como cada recurso influencia a saída do modelo. Isso é particularmente útil para entender a importância relativa de diferentes características na tomada de decisões do modelo.

A outra técnica utilizada para prover explicações globais são os Gráficos de Dependência Parcial, PDP (*Partial Dependence Plots*), que ilustram a relação entre a resposta desejada e características específicas de entrada, levando em consideração os valores de outras características. Devido às limitações da percepção humana, os PDPs geralmente se concentram em uma ou duas características-chave para mostrar a resposta desejada esperada.

C. *Contrafactual*

O DiCE é uma ferramenta que gera explicações contrafactuais para modelos de aprendizado de máquina, mostrando versões de determinadas amostras com características adicionadas de perturbações. Por exemplo, gerar uma amostra de uma pessoa que teria recebido um empréstimo a partir de uma amostra à qual o empréstimo foi negado. Ele fornece explicações do tipo “e se” para a saída do modelo, sendo útil tanto para usuários finais quanto para desenvolvedores de modelos. O DiCE consegue gerar explicações contrafactuais para qualquer modelo de aprendizado de máquina tratando o problema como uma otimização. Ele é capaz de gerar um conjunto de explicações contrafactuais ajustáveis para diversidade e proximidade em relação à entrada original, além de suportar restrições simples para garantir a viabilidade dos exemplos gerados.

IV. RESULTADOS

Para os modelos interpretáveis diretamente, realizamos o treinamento de cada modelo utilizando uma aumentação do conjunto de dados, encontrando o modelo com menor Erro absoluto médio. Apesar das métricas de regressão serem importantes para o modelo, não limitamos nossa análise sobre esta métrica e, portanto, realizamos uma análise mais aprofundada dos modelos propostos. Cada modelo foi submetido a uma validação cruzada com 10 *folds* e reportamos as métricas de **Erro absoluto médio**, para uma análise de performance geral do modelo, Acurácia Balanceada no rótulo de maior frequência, para medir a capacidade de distinção do modelo em áreas com menor índice de violência, com o objetivo de melhorar a performance geral do regressor. Também realizamos análises através de uma métrica de consistência para regressão, para verificar possíveis vieses do modelo.

$$\text{consistency error} = \mathbb{E}_{x \sim X} [f_{\theta}(x) - \mathbb{E}_{x' \sim \mathcal{N}_k(x)} [f_{\theta}(x')]]$$

Onde $\mathcal{N}_k(x)$ indica os k vizinhos mais próximos de x . Os resultados, apresentados na Tabela I e, apesar dos resultados com alto erro absoluto médio, não nos preocupamos tanto com esta métrica, pois nosso interesse não está em prever com exatidão a quantidade de crimes, mas extrair um modelo capaz de extrair explicações para o problema de violência. Deste modo, consideramos a acurácia balanceada como característica mais importante e reportamos na tabela a acurácia em prever zeros.

Pode-se observar que os modelos em suas versões *Zero-inflated* tiveram resultados superiores em métricas de performance em relação a sua contra-parte padrão, exceto pela métrica de fairness, muito provavelmente devido a introdução de uma rede de classificação separada, que pode estar induzindo a novos vieses ao filtrar as amostras que passam para o regressor. Apesar dessa característica resolvemos realizar uma análise mais detalhada dos modelos *Zero-inflated* através da visualização das predições e dos resíduos, conforme pode ser visto em Apêndice, que inclui análises de acurácia balanceada para diferentes cortes. Com isso, julgamos que o modelo de Poisson *Zero-inflated*, apesar de possuir alguns erros sistemáticos, ainda atingia uma performance superior quando analisada em sua totalidade e era apropriado interpretá-lo.

O modelo *Zero-inflated* utiliza de dois modelos, um para diferenciar locais com rótulo zero dos rótulos maiores que zero e outro para prever a quantidade de boletins de ocorrência. Como o classificador utilizado foi um modelo de regressão logístico e, portanto, diretamente interpretável e a regressão de Poisson é um modelo linear generalizado, com função de *link* sendo logarítmica ao invés da logit, como é realizada na regressão logística, o regressor também é interpretável. Os pesos ω de uma regressão de Poisson podem ser diretamente comparados com a saída do modelo pela relação

$$\mathbb{E}[y|x] = \exp(\omega^T x).$$

Então, o modelo escolhido pode prover duas explicações diferentes, uma para cada modelo, que estão mostradas na Figura 1. É possível visualizar alguns padrões que aparecem em ambos os modelos, incluindo as features que mais determinam alta criminalidade: Muito alta densidade e Áreas Rurais.

Acreditamos que estas características sejam mais comuns em aspectos distintos de crimes e, como não realizamos qualquer distinção dentre crimes, por exemplo, entre assaltos e roubo de carros, as características tendem a refletir um padrão de crime mais geral. A feature referente a áreas rurais pode refletir mais roubo de carros, que possuía uma base de dados só para esta modalidade. Por sua vez, a feature de alta densidade pode refletir mais os crimes urbanos, que podem realmente estar relacionados com a alta densidade demográfica, característica comum de áreas com população carente.

Tabela I
MÉTRICAS OBTIDAS POR VALIDAÇÃO CRUZADA DOS MODELOS PROPOSTOS.

	Model	MAE	Bal Acc (%)	Consistency error
Standard	Linear Regressor	46.008 ± 1.041	63.654 ± 1.015	16.852 ± 0.163
	Poisson Regressor	40.102 ± 0.853	50.355 ± 0.082	14.733 ± 0.275
	Quantile Regressor	40.898 ± 1.555	64.721 ± 0.990	5.881 ± 0.062
Zero-Inflated	Linear Regressor	45.333 ± 0.847	67.850 ± 0.896	17.656 ± 0.190
	Poisson Regressor	40.048 ± 0.838	57.222 ± 0.456	14.859 ± 0.269
	Quantile Regressor	40.346 ± 1.560	67.825 ± 0.789	6.193 ± 0.057

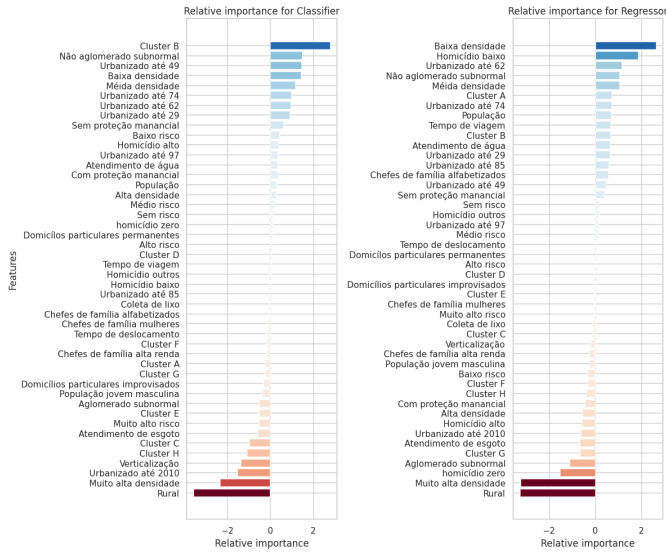


Figura 1. Importância relativa para cada modelo. Quanto mais positiva a importância, mais ela contribui para a queda no rótulo.

Outra característica importante que podem ser visualizadas são a de urbanização até a década de 2010, que se correlaciona com a expansão urbana da cidade e construção de casas populares destinadas a uma população carente, que ainda sofre com a criminalidade. No entanto, a feature de Homicídio zero, que foi considerada como marcadora de altas taxas criminais foi inesperado. Apesar do rótulo de criminalidade não conter registro de homicídios, apenas de roubos, esperávamos ver alguma correlação entre as taxas de homicídio e roubo numa mesma região, no entanto, os resultados destes modelos não nos diz isso. A hipótese que temos é a de sub-notificação de homicídios em diversas regiões de São Paulo, o que dificulta na associação das duas características.

Outras features interessantes, agora para caracterizar regiões com menos registros de boletins de ocorrência foram a Baixa densidade demográfica, Urbanizações mais antigas, Regiões que não aglomerados subnormais e regiões no Cluster B, ou seja, regiões parecidas com o centro expandido de São Paulo.

Para os modelos não-interpretáveis, realizamos a mesma rotina de treinamento dos demais modelos, utilizando dados aumentados e validando por validação cruzada. Os resultados destes modelos, na Tabela II, foi utilizado para selecionar a regressão Quantílica, que obteve resultados consideravelmente superiores, para ser explicada.

Tabela II
MÉTRICAS OBTIDAS POR VALIDAÇÃO CRUZADA DOS MODELOS NÃO-INTERPRETÁVEIS.

Model	MAE	Bal Acc	Consistency error
Standard	29.904 ± 3.279	56.075 ± 0.857	43.858 ± 1.376
Poisson	26.889 ± 3.543	50.532 ± 0.173	42.858 ± 0.766
Quantile	33.806 ± 1.472	65.263 ± 2.296	23.757 ± 0.453

O primeiro método de explicação explorado foi o SHAP, que permite uma visualização da contribuição das features para diferentes amostras. Utilizando o modelo de Árvore de decisão por Gradient Boosting Quantílico, realizamos uma visualização em conjunto de teste, conforme a Figura 2.

A visualização gerada pelo método nos ajuda a tirar conclusões em sua importância, principalmente como os diferentes valores para a feature impacta a importância dela no modelo. Por exemplo, as features que mais impactam a saída do modelo são aquelas no topo da Figura 2, A densidade demográfica, o total de viagens, a População e taxas de homicídio.

Para a gerar uma explicação baseada em PDP, utilizamos o mesmo modelo empregado no SHAP. Podemos observar, nos gráficos presentes no Apêndice, que a proporção de chefes de família alfabetizados tem um impacto significativo na quantidade de BOs de determinadas regiões, e quanto maior essa proporção, menor a criminalidade. Outras features que demonstraram uma correlação com a periculosidade das regiões foram a densidade populacional, a verticalização, a proporção de população jovem masculina e o número total de viagens. É interessante notar que estes resultados contribuem para a explicação do SHAP, já que ambos concordam na forma que a mudança nas características impacta na saída do modelo, por exemplo, na Densidade demográfica, que quanto maior, mais impacta negativamente na saída.

Outra explicação interessante que surge na interpretação do modelo mais simples é do impacto no Cluster para a criminalidade. Na explicação para Cluster no PDP, verificamos que dois cluster têm um maior impacto positivo, estes sendo A e B, regiões centrais de São Paulo e, possivelmente, com melhor policiamento.

Já para gerar explicação contrafactual, empregamos a biblioteca DiCE, tendo como *target* a quantidade de 'BO'. Nesse contexto, definimos como 20 o número máximo de BOs. O treinamento foi realizado também utilizando o modelo HistGradientBoostingRegressor. Na Tabela IV, são exibidos exemplos de 3 contrafactuais gerados, onde é possível observar que os atributos PJM10, DEN10, CAR10 e CAL10 precisa-

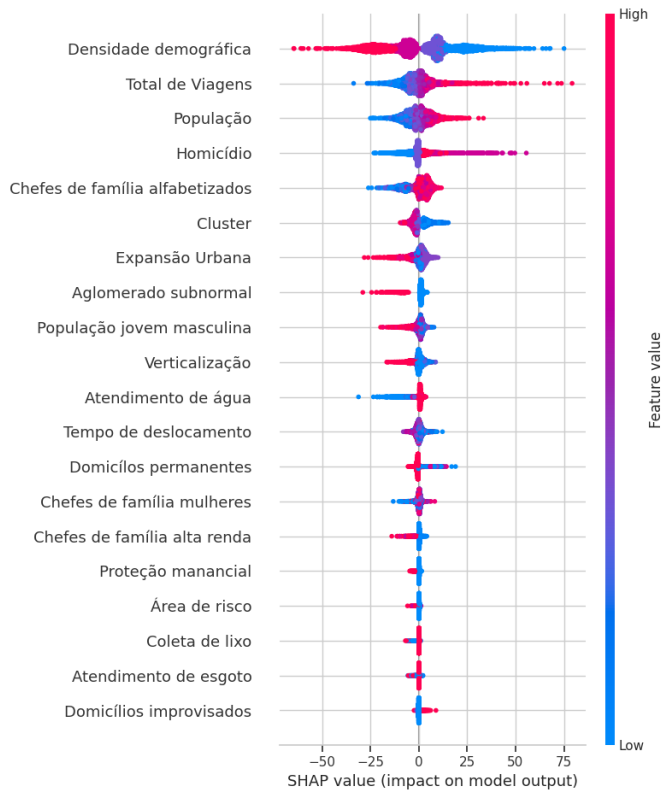


Figura 2. Explicação SHAP no conjunto de teste

riam ser modificados na amostra testada para que a predição atual de 245 BOs seja reduzida para menos de 20. Podemos analisar, também, que em regiões com uma alta incidência de BOs, há uma consistente indicação de que uma mudança é necessária no atributo DEN10.

Tabela III
AMOSTRA DE CONSULTA (SAÍDA ORIGINAL : 245.0)

PJM10	DEN10	CAR10	CAL10	BO
0.0807	1	0.0	0.9527	245.0

Tabela IV
EXEMPLO DE CONTRAFATUAIS (NOVA SAÍDA: [0, 20])

PJM10	DEN10	CAR10	CAL10	BO
-	4.0	0.4	-	11.33
-	4.0	-	0.6	15.46
0.4	4.0	-	-	13.14

V. DISCUSSÃO

Neste trabalho exploramos técnicas de interpretabilidade em modelos de aprendizado de máquina, onde optamos por uma abordagem diferente dos trabalhos anteriores, modelando uma regressão para prever a distribuição dos boletins de ocorrência.

É interessante também destacar a concordância dos métodos em prover uma explicação para o problema de criminalidade,

destacando aspectos sociais, de acesso a moradia, educação e transporte como fatores que impactam na segurança e bem-estar de uma população que sofre com a criminalidade.

Nossos resultados estão de acordo com o esperado levando em consideração as discussões sobre desigualdade social que vêm ocorrendo atualmente. Utilizando o SHAP, PDP e encontrando os contrafactuais, verificamos que as características que mais impactam na distribuição dos BOs são a renda familiar, a densidade populacional, a distância de determinadas regiões do centro da cidade e o nível de alfabetização da população.

A partir destes resultados, e ao propor um modelo de aprendizado de máquina para entender essas raízes, concluímos na importância de visualizar o processo de decisão do modelo. As diferentes conclusões obtidas entre o modelo interpretável, que capturava de forma mais superficial, para o modelo complexo, munido de métodos de explicabilidade mais refinados, mostram que a explicabilidade em aprendizado de máquina é importante e acessível. Sendo a explicação em vezes mais importante que apenas a saída do modelo, com ênfase para decisões com alto impacto, como alocação de recursos para regiões mais carentes, mas principalmente como alocar tais recursos.

REFERÊNCIAS

- [1] Xu Zhang, Lin Liu, Minxuan Lan, Guangwen Song, Luzi Xiao, and Jinguo Chen. Interpretable machine learning models for crime prediction. *Computers, Environment and Urban Systems*, 94:101789, 2022.
- [2] Andrew P Wheeler and Wouter Steenbeek. Mapping the risk terrain for crime using machine learning. *Journal of Quantitative Criminology*, 37:445–480, 2021.
- [3] Sunjae Kim and Sugie Lee. Nonlinear relationships and interaction effects of an urban environment on crime incidence: Application of urban big data and an interpretable machine learning method. *Sustainable Cities and Society*, 91:104419, 2023.
- [4] James M Hickey, Pietro G Di Stefano, and Vlasios Vasileiou. Fairness by explicability and adversarial shap learning. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2020, Ghent, Belgium, September 14–18, 2020, Proceedings, Part III*, pages 174–190. Springer, 2021.
- [5] Rafael Prieto Curiel, Sofia Collignon Delmar, and Steven Richard Bishop. Measuring the distribution of crime and its concentration. *Journal of quantitative criminology*, 34:775–803, 2018.

APÊNDICE

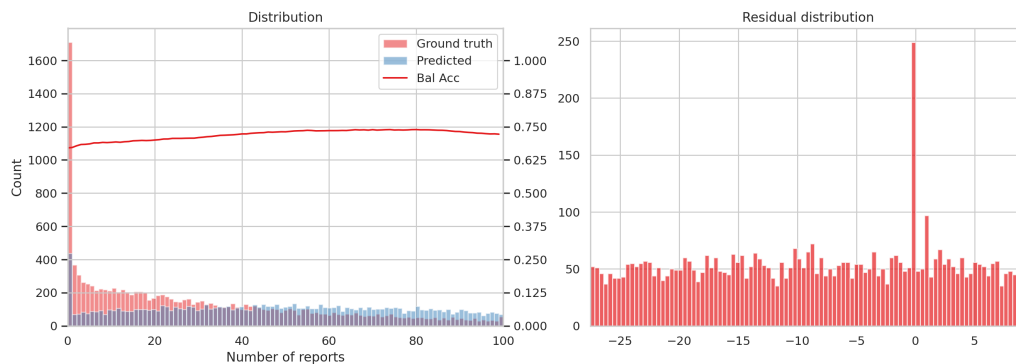


Figura 3. Distribuições do rótulo previsto pelo modelo de Regressão Linear *Zero-inflated*. É possível ver que os resíduos se distribuem uniformemente, errando com mais frequência para baixo. Essa característica não é uma boa qualidade e não demonstra confiança na predição, portanto, descartamos este modelo como preditor.

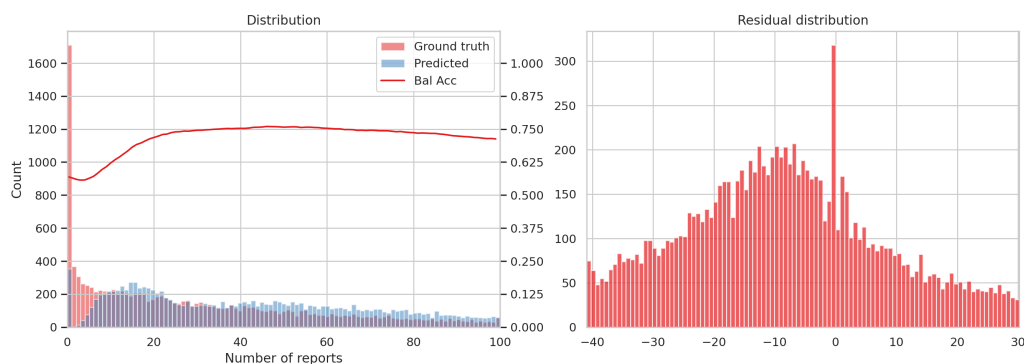


Figura 4. Distribuições das previsões do modelo de Poisson *Zero-inflated*. Os resíduos aqui aparecem centrados em -10, indicando um erro sistemático através de um possível viés de modelagem e uma correção, através de outras técnicas, pode melhorar ainda mais o modelo. A acurácia balanceada se encontra bem distribuída, indicando um modelo bem ajustado.

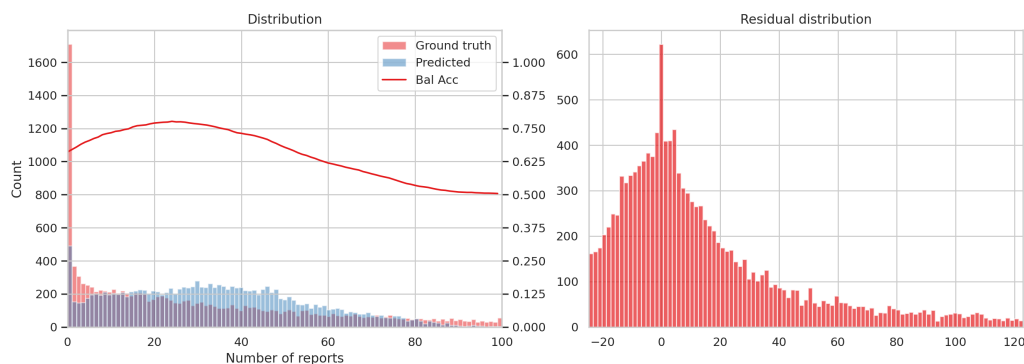


Figura 5. Distribuição da Regressão Quantílica *Zero-inflated*. O modelo obteve boa performance para prever rótulos com rótulos menores, trazendo toda a distribuição da cauda direita para o intervalo entre 20 e 80. Apesar desta estratégia ter centralizado ainda mais os resíduos para perto de zero, o modelo ainda atinge uma acurácia parecida com outros modelos, piorando conforme mais avançamos no valor do rótulo.

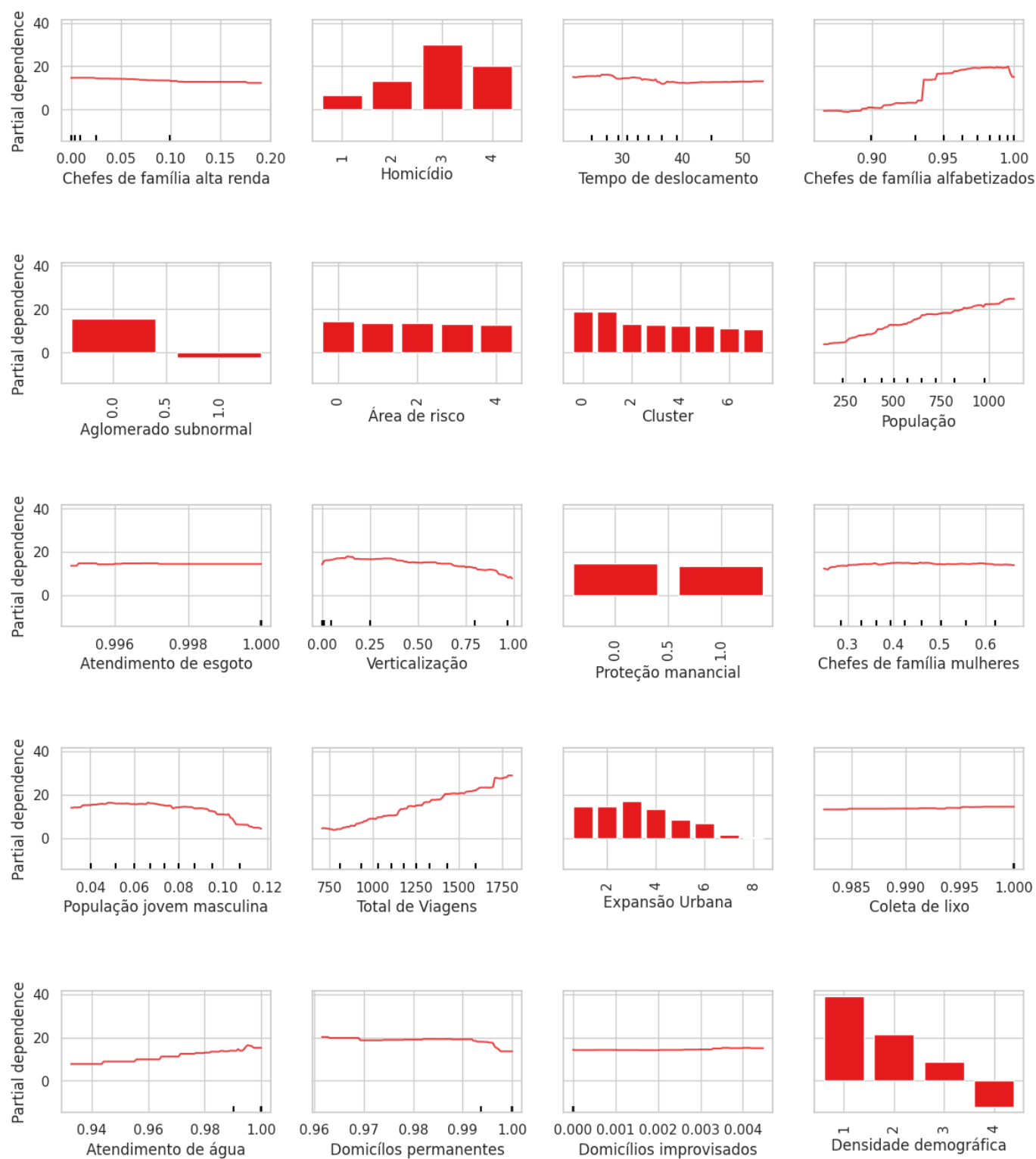


Figura 6. Dependências parciais de cada feature individualmente.