UNIVERSIDADE ESTADUAL PAULISTA - UNESP
"JÚLIO DE MESQUITA FILHO"
Faculdade de Ciências e Tecnologia - Campus de Presidente Prudente

ANDERSON GARRIDO SCAIONI

Policiamento Inteligente: Aplicação de Ontologia, Análise Preditiva e Redes Neurais Convolucionais no Apoio ao Planejamento Operacional e Roteirização

Presidente Prudente
2025

ANDERSON GARRIDO SCAIONI

Policiamento Inteligente: Aplicação de Ontologia, Análise Preditiva e Redes Neurais Convolucionais no Apoio ao Planejamento Operacional e Roteirização

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Faculdade de Ciências e Tecnologia da Universidade Estadual Paulista UNESP, como requisito para a obtenção do título de Mestre em Ciência da Computação.

Presidente Prudente
2025

S287p

Scaioni, Anderson Garrido
Policiamento Inteligente : Aplicação de Ontologia, Análise Preditiva e Redes Neurais Convolucionais no Apoio ao Planejamento Operacional e Roteirização / Anderson Garrido Scaioni. -- Presidente Prudente, 2025
131 p. : il., tabs., mapas

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Tecnologia, Presidente Prudente
Orientador: Rogério Eduardo Garcia

1. Policiamento inteligente. 2. Predição Criminal. 3. Redes Neurais. 4. Ontologia. 5. Segurança Pública. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Dados fornecidos pelo autor(a).

IMPACTO POTENCIAL DESTA PESQUISA

A pesquisa contribui com inovação na segurança pública, ao aplicar ontologia e inteligência artificial no planejamento policial, promovendo eficiência operacional, uso racional de recursos e benefícios sociais.

POTENTIAL IMPACT OF THIS RESEARCH

This research contributes with innovation in public safety by applying ontology and artificial intelligence to police planning, promoting operational efficiency, rational resource use, and social benefits.

UNIVERSIDADE ESTADUAL PAULISTA

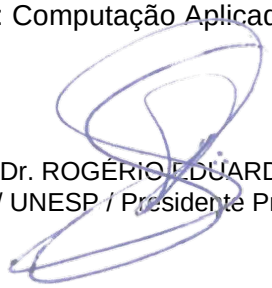
Câmpus de Presidente Prudente

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Emprego de Redes Neurais Convolucionais, Ontologia e Análise Preditiva para o planejamento estratégico e roteirização eficiente, como proposta de policiamento inteligente à Segurança Pública

AUTOR: ANDERSON GARRIDO SCAIONI
ORIENTADOR: ROGÉRIO EDUARDO GARCIA

Aprovado como parte das exigências para obtenção do Título de Mestre em Ciência da Computação, área: Computação Aplicada pela Comissão Examinadora:



Prof. Dr. ROGÉRIO EDUARDO GARCIA (Participaçao Presencial)
FCT / UNESP / Presidente Prudente (SP)

Prof. Dr. DANILLO ROBERTO PEREIRA (Participaçao Presencial)
Matemática e Computação / Unesp - Presidente Prudente - FCT

Prof. Dr. WASHINGTON HENNIS DA SILVA (Participaçao Presencial)
18 BPM/1 / PMESP

Presidente Prudente, 26 de fevereiro de 2025

Dedico este trabalho aos meus pais, que sempre me incentivaram, e àqueles que acreditaram no poder da educação como agente transformador.

Agradecimentos

A Deus, por me dar forças e sabedoria para seguir em frente.

Aos meus familiares, pelo apoio incondicional em todos os momentos desta caminhada.

Ao meu orientador, Prof. Dr. Rogério Eduardo Garcia, pela orientação segura, paciência, dedicação e por acreditar na proposta deste trabalho desde o início. Sua expertise, comprometimento e incentivo foram fundamentais para a construção e conclusão desta dissertação.

Aos professores e colegas do Programa de Pós-Graduação em Ciência da Computação da UNESP, pelo compartilhamento de conhecimento e convivência ao longo dessa jornada.

À Polícia Militar do Estado de São Paulo, pelo apoio institucional.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho, minha sincera gratidão.

"Não sabendo que era impossível, foi lá e fez."
Mark Twain

Resumo

Este trabalho implementa um protótipo de um sistema de policiamento inteligente utilizando tecnologias como Redes Neurais Convolucionais, Ontologia e Análise Preditiva. O foco é otimizar o planejamento operacional, fornecendo ferramentas para o roteamento eficiente de recursos dentro da Polícia Militar, garantindo uma melhor cobertura e resposta aos incidentes. O sistema de análise preditiva identificará padrões e tendências, enquanto experimentos com dados do mundo real avaliarão a eficácia das abordagens tecnológicas propostas. Modelos estabelecidos de ciência de dados, incluindo redes neurais convolucionais treinadas, serão aplicados e testados dentro de estruturas de treinamento. As autoridades públicas necessitam de ferramentas práticas para combater o crime, incluindo aquelas que auxiliem na prevenção de novas ocorrências criminosas, otimização da alocação de recursos, estudo de comportamentos e padrões criminosos, e visualização de áreas geográficas de alta criminalidade. O objetivo principal é aprimorar a eficiência das Unidades Policiais na prevenção e combate ao crime, por meio de processos de tomada de decisão e alocação de recursos mais eficazes. Os resultados esperados incluem o desenvolvimento de um modelo integrado de policiamento inteligente, contribuindo para a redução das taxas de criminalidade e para melhores práticas de gestão na segurança pública. Sendo assim, este estudo busca fornecer uma abordagem precisa e eficiente para apoiar as decisões sobre o roteamento de patrulhas, aumentando a eficácia das respostas e melhorando a segurança pública de maneira geral.

Palavras-chave: Policiamento inteligente, Redes Neurais Convolucionais, Ontologia, Análise Preditiva, Planejamento Operacional, Roteamento de Recursos, Prevenção ao Crime, Alocação de Recursos, Padrões Criminosos, Visualização Geográfica, Ciência de Dados, Segurança Pública, Tomada de Decisão, Modelos Preditivos, Patrulhamento Otimizado.

Abstract

This work implements a prototype of an intelligent policing system using technologies such as Convolutional Neural Networks, Ontology, and Predictive Analysis. The focus is to optimize operational planning by providing tools for the efficient routing of resources within the Military Police, ensuring better coverage and response to incidents. The predictive analysis system will identify patterns and trends, while experiments using real-world data will evaluate the effectiveness of the proposed technological approaches. Established data science models, including trained convolutional neural networks, will be applied and tested within training frameworks. Public authorities require practical tools to combat crime, including those that assist in preventing new criminal occurrences, optimizing resource allocation, studying criminal behaviors and patterns, and visualizing high-crime geographic areas. The main objective is to enhance the efficiency of Police Units in crime prevention and response through more effective decision-making and resource allocation processes. Expected outcomes include the development of an integrated intelligent policing model, contributing to lower crime rates and improved public safety management practices. Thus, this study aims to provide a precise and efficient approach to support decisions on patrol routing, increasing response effectiveness and improving overall public safety.

Keywords: Intelligent Policing, Convolutional Neural Networks, Ontology, Predictive Analysis, Operational Planning, Resource Routing, Crime Prevention, Resource Allocation, Criminal Patterns, Geographic Visualization, Data Science, Public Security, Decision-Making, Predictive Models, Optimized Patrolling.

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

ABNT	Associação Brasileira de Normas Técnicas
CNN	Convolutional Neural Network (Rede Neural Convolucional)
CSV	Comma-Separated Values (Valores Separados por Vírgula)
FCT	Faculdade de Ciências e Tecnologia
GESPOL	Sistema de Gestão da Polícia Militar do Estado de São Paulo
LDO	Lei de Diretrizes Orçamentárias
LOA	Lei Orçamentária Anual
MAE	Mean Absolute Error (Erro Absoluto Médio)
PMESP	Polícia Militar do Estado de São Paulo
PPA	Plano Plurianual
PPGCC	Programa de Pós-Graduação em Ciência da Computação
RDF	Resource Description Framework
RTM	Risk Terrain Modeling (Modelagem de Terreno de Risco)
SP	São Paulo
UNESP	Universidade Estadual Paulista

Lista de Símbolos

α	Taxa de aprendizado (learning rate)
ϵ	Erro admissível ou limiar de convergência
θ	Parâmetro de peso em rede neural
σ	Função de ativação sigmoide
μ	Média
σ^2	Variância
X	Conjunto de variáveis independentes (features)
Y	Variável dependente ou variável de saída
$f(x)$	Função preditiva
$\hat{y}$	Valor estimado (previsão)
MAE	Mean Absolute Error (Erro Absoluto Médio)
MSE	Mean Squared Error (Erro Quadrático Médio)
$RMSE$	Raiz do Erro Quadrático Médio
k	Número de clusters (em algoritmos de agrupamento)
n	Número total de observações

Sumário

CAPÍTULO

1

Proposta do Projeto

1.1 Considerações Iniciais

A Segurança Pública é um tema de grande importância para a sociedade, e aprimorar as medidas de gestão nessa área é fundamental para garantir a proteção dos cidadãos. Nesse contexto, o policiamento inteligente surge como uma alternativa inovadora e necessária, especialmente quando a quantidade de dados disponíveis é cada vez maior, mas a forma de planejamento permanece a mesma há décadas. As forças policiais enfrentam diversos desafios no desempenho de suas funções, como a falta de informações precisas e em tempo hábil, a dificuldade na tomada de decisões e a alta demanda por recursos humanos e materiais. Nesse contexto, a Ciência da Computação pode contribuir para aprimorar a eficiência da Polícia Militar, permitindo uma melhor compreensão do contexto em que as ocorrências são registradas e possibilitando a identificação de padrões e tendências.

Esta dissertação propõe o emprego de Redes Neurais Convolucionais, Ontologia e Análise Preditiva para o planejamento estratégico e roteirização eficiente, como proposta de policiamento inteligente à Segurança Pública. O foco é contribuir para a melhoria da eficiência da Polícia Militar na prevenção e combate ao crime, através do aprimoramento dos processos de tomada de decisão e alocação de recursos.

A recepção das ocorrências no âmbito da Segurança Pública é um processo crucial que impacta diretamente a eficácia das ações policiais. Atualmente, essas ocorrências são recebidas no formato JSON, o que reflete uma evolução na padronização e estruturação dos dados. Apesar dessa organização mais consistente, a interpretação eficiente desses registros ainda apresenta desafios significativos. O formato variado dos registros, provenientes de dife-

rentes fontes, pode dificultar a análise e a compreensão sistêmica dos eventos. Nesse contexto, técnicas avançadas de normalização e mapeamento desempenham um papel essencial na tese proposta. A normalização visa estabelecer um padrão consistente para os dados, facilitando a comparação e a análise, enquanto o mapeamento permite a associação eficiente de informações de diferentes fontes, contribuindo para uma visão integrada do cenário. Essas técnicas serão aplicadas de forma integrada à proposta de policiamento inteligente, utilizando Redes Neurais Convolucionais, Ontologia e Análise Preditiva. O objetivo é não apenas aprimorar a eficiência operacional da Polícia Militar, mas também otimizar a assimilação e interpretação das ocorrências em formato JSON, tornando-as mais acessíveis para embasar decisões estratégicas e alocação de recursos.

Com base nessa contextualização, o presente trabalho busca propor uma abordagem baseada em ontologias e análise preditiva para aprimorar a eficiência da Polícia Militar na prevenção e combate à criminalidade, por meio da utilização de dados históricos e em tempo real para identificação de padrões e tendências. Métodos de Policiamento Preditivo (*Predictive Policing*) e Policiamento Inteligente (*Smart Policing*) têm sido desenvolvidos com o objetivo de reduzir as ocorrências criminais. As propostas existentes se destacam por priorizarem ações proativas voltadas à antecipação de eventos criminais, baseadas na utilização de dados com análise aprimorada, acompanhamento de desempenho e incentivo à inovação para o emprego e desenvolvimento de estratégias e táticas de policiamento adequadas (Coldren Jr et al., 2013, Perry, 2013, Rummens et al., 2017).

A partir dos conceitos de *Smart Policing* e de *Predictive Policing*, superficialmente discutidos aqui, a proposta de implementação de um sistema de análise preditiva para identificar padrões e tendências, usando dados históricos e em tempo real, é apresentada. Os resultados incluem o desenvolvimento de um modelo integrado de policiamento inteligente e a contribuição para a redução dos índices criminais, além do aprimoramento das medidas de gestão na área de Segurança Pública.

Cabe salientar que, não havendo segurança, as pessoas alteram suas rotinas de trabalho e lazer, gerando medo e ansiedade. Quando as pessoas transferem parte da sua atenção, preocupadas em viabilizar suas atividades, visando à garantia de sua integridade física, elas prejudicam sua qualidade de vida e isso afeta o bem-estar geral da sociedade (Grabosky, 1995, Basilio et al., 2019, Baig et al., 2024). A partir de uma perspectiva de bem-estar social da coletividade, a segurança pública se apresenta como uma questão fundamental.

Como consequência, há que se haver um aperfeiçoamento da gestão, es-

pecialmente da administração pública. Nesse contexto, são exigidas eficiência e eficácia na efetividade, uma vez que as ações prestacionais da administração pública são custeadas pelos cidadãos-contribuintes e, consequentemente, a população demanda qualidade nas ações de governo (Carvalho e Tonet, 1994)(p. 139). Logo, a tomada de decisão para gestores públicos deve buscar otimizar recursos e resultados, para atender o interesse público, e compete à administração avaliar e escolher, no plano concreto, as melhores soluções (Freitas, 2008)

Nesse cenário, tem sido apontada a necessidade de aperfeiçoar o processo de gerenciamento dos órgãos de segurança, desenvolver tecnologias para a atividade policial e inovar nos processos de tomada de decisão e planejamento, para melhorar a segurança pública, bem como promover maior integração com a sociedade (Beato Filho, 2009, Ballesteros, 2014). Percebe-se que o sucesso das ações depende não apenas do estabelecimento de um sistema de gestão estratégica, mas também da utilização de ferramentas para tomada de decisão correta, o que impacta diretamente na redução dos crimes. Algumas mudanças nas atividades policiais não decorrem de soluções estruturais em larga escala, mas sim de alterações no modo habitual de trabalho dentro das instituições, conforme apontado por Beato Filho (2009).

Apesar do aumento no número de pesquisas na área, estudos sobre policiamento inteligente e preditivo ainda são escassos. De acordo com pesquisas de revisão bibliográfica, embora modelos de otimização na segurança pública tenham sido aplicados há décadas, os primeiros trabalhos com ênfase na predição datam de 2010[(Meijer e Wessels, 2019, Ratcliffe, 2010).

1.2 Problema, Questões de pesquisa, Hipótese e tese

Apesar do crescente volume de dados disponíveis, os métodos de planejamento e execução do policiamento ostensivo preventivo em locais de risco permanecem relativamente inalterados há décadas. Diante da necessidade de aprimoramento do planejamento das ações de policiamento preventivo, é possível prevenir o crime, ao invés de reagir à sua ocorrência. É possível identificar padrões, gerar subsídios para tomada de decisões de gestores de polícia, tornando policiamento eficaz e eficiente.

Diante da complexidade e interdisciplinaridade da tecnologia de aplicação de Redes Neurais Convolucionais (CNNs), Ontologia e Análise Preditiva no contexto do policiamento inteligente, surgem questões pertinentes. São elas:

1. Quais desafios éticos emergem da implementação da análise preditiva no policiamento e como podem ser mitigados?

2. Como as CNNs podem ser otimizadas para identificar padrões comportamentais específicos em ambientes urbanos?

3. De que maneira a ontologia pode ser estruturada para garantir uma representação semântica abrangente dos dados de segurança pública?

Estas questões fundamentam a investigação, orientando a busca por soluções inovadoras e sustentáveis que maximizem o potencial dessas tecnologias na promoção da segurança pública eficaz e responsável.

No contexto da interseção entre Redes Neurais Convolucionais (CNNs), Ontologia e Análise Preditiva para o aprimoramento do policiamento inteligente, a formulação de hipóteses é essencial para guiar a pesquisa.

A tese subjacente a este estudo é a de que a implementação integrada de Redes Neurais Convolucionais (CNNs), Ontologia e Análise Preditiva oferece uma abordagem transformadora no âmbito do policiamento inteligente, promovendo uma gestão mais eficiente e proativa da segurança pública. Propomos que, ao incorporar essas tecnologias de forma sinérgica, é possível criar um ambiente inovador capaz de antecipar, prevenir e responder de maneira mais eficaz a atividades criminosas.

A tese baseia-se na premissa de que as CNNs, ao processarem informações visuais, podem identificar padrões complexos e comportamentos suspeitos, fortalecendo a capacidade de vigilância e monitoramento urbano. A integração de uma ontologia robusta visa proporcionar uma compreensão semântica aprimorada dos dados, facilitando a interoperabilidade entre diferentes fontes de informação e melhorando a tomada de decisões.

Ademais, a Análise Preditiva emprega algoritmos avançados, o que possibilita não apenas a resposta reativa a eventos, mas também a antecipação de cenários, permitindo um planejamento estratégico mais eficiente. Embora fora do escopo do mestrado [1], há evidência da aplicação e de vantagens com o uso de dados históricos. Desse modo, considera-se suficiente para considerar que a abordagem proposta, quando aplicada ao policiamento, tem o potencial de gerar resultados tangíveis na redução da criminalidade, na otimização do uso de recursos e na construção de comunidades mais seguras e resilientes. Evidências e análises aprofundadas são apresentadas de modo a validar e sustentar essa tese, contribuindo para o avanço do conhecimento na convergência entre tecnologias emergentes e segurança pública.

1.3 Metodologia

Com mote no problema de pesquisa proposto, a dissertação irá considerar os seguintes objetivos:

[1] Considera-se fora do escopo porque a aplicação em um ambiente real para mostrar sua aplicação e redução da criminalidade está fora do alcance, uma vez que depende de autorização e concordância da Polícia Militar do Estado de São Paulo

1.3.1 Objetivo Geral

Propor um modelo de policiamento inteligente integrado utilizando redes neurais e análise preditiva para propor a localização preventiva de patrulhas policiais e sua roteirização, para aprimorar a tomada de decisão de gestores e, por consequência, melhorar a eficiência da Polícia Militar na prevenção e combate à criminalidade, por meio da utilização de dados históricos.

1.3.2 Objetivos Específicos

- Desenvolver um modelo de análise preditiva baseado em redes neurais convolucionais, que utilize dados históricos para identificar padrões e tendências de criminalidade em determinadas áreas geográficas.
- Integrar o modelo de análise preditiva com padrões ontológicos para fornecer uma visão mais completa e precisa da criminalidade em determinadas áreas geográficas, permitindo que os gestores da Polícia Militar tomem decisões informadas sobre o policiamento.
- Verificar a possibilidade de emprego de método de análise preditivo para os crimes definidos como passíveis de bonificação no Estado de São Paulo, de acordo com a área dos municípios em estudo.
- Identificar e avaliar de fatores de risco, ocorrências correlatas e ambientais, situações climáticas e econômicas que envolvem os crimes como por exemplo: furtos, perturbações do sossego público, desinteligências, dentre outros na região central da Capital Paulista.
- Analisar a possibilidade de disposição de ferramenta integrada de predição e otimização nas atividades de policiamento preventivo.

1.4 Fundamentação Teórica

A fundamentação teórica deste projeto envolve conceitos essenciais da Ciência da Computação aplicados à Segurança Pública, com destaque para as Redes Neurais Convolucionais, a Ontologia, a Análise Preditiva e o conceito de Policiamento Inteligente, os quais sustentam a proposta metodológica deste estudo.

As Redes Neurais Convolucionais (CNNs) exercem um papel fundamental no tratamento de dados espaciais, destacando-se como um dos principais paradigmas do aprendizado profundo. Inspiradas na estrutura do cérebro humano, essas redes são particularmente eficazes na detecção e identificação de padrões em imagens e mapas geográficos. Sua capacidade de reconhecer estruturas hierárquicas e complexas as torna ideais para o processo de análise georreferenciada, contribuindo diretamente para a interpretação de dados e

para a tomada de decisão em ambientes com alta densidade de informação visual.

A Ontologia, por sua vez, é empregada neste projeto como uma forma de representar o conhecimento de maneira estruturada e semanticamente enriquecida. Através da modelagem ontológica, é possível organizar hierarquicamente conceitos relacionados à Segurança Pública e estabelecer relações entre variáveis críticas. Essa representação semântica amplia a capacidade de análise dos dados, facilitando tanto a busca por informações quanto a compreensão de interdependências complexas entre ocorrências, territórios e padrões criminais.

A Análise Preditiva é outro eixo central da proposta, sendo utilizada para identificar padrões históricos e tendências em bases de dados criminais. Por meio da aplicação de técnicas estatísticas e algoritmos de aprendizado de máquina, essa abordagem permite prever com maior precisão áreas e períodos de maior risco, orientando o planejamento operacional da Polícia Militar de forma proativa. Ao antecipar cenários críticos, torna-se possível otimizar a alocação de recursos e adotar medidas preventivas mais eficazes.

O conceito de Policiamento Inteligente representa uma evolução das práticas tradicionais de segurança pública. Essa abordagem está baseada na utilização integrada de tecnologias avançadas e dados em tempo real para aprimorar a tomada de decisões estratégicas. Neste projeto, propõe-se a integração entre redes neurais convolucionais, ontologias e análise preditiva, a fim de compor um modelo sistêmico de apoio ao planejamento e à operação policial. O objetivo é proporcionar respostas mais rápidas e eficientes, potencializando a prevenção de crimes e promovendo melhorias significativas na gestão da segurança pública.

1.5 Metodologia Aplicada

Este trabalho será conduzido por meio de uma pesquisa qualitativa e exploratória, com base em dados empíricos obtidos por meio de revisão bibliográfica, entrevistas e visitas técnicas. A metodologia inclui a implementação de um sistema de análise preditiva baseado em redes neurais convolucionais, voltado à identificação de padrões e tendências a partir de dados históricos e em tempo real. Os experimentos serão realizados com dados reais, com o intuito de avaliar a eficácia da abordagem proposta.

A seguir, descrevem-se as principais etapas da pesquisa:

1.5.1 Definição e Contextualização da Área de Estudo

Inicialmente, será realizada a seleção e delimitação da área geográfica onde o modelo de policiamento inteligente será aplicado. Esta fase inclui a análise de dados demográficos, criminais e geoespaciais da região selecionada, com o

intuito de compreender o ambiente operacional da Polícia Militar.

Nesta etapa, a função de sazonalidade será considerada como critério preditivo fundamental. Eventos sazonais, cíclicos ou irregulares como férias escolares, feriados prolongados, finais de semana e variações climáticas impactam diretamente os índices criminais e serão incorporados aos modelos analíticos. Compreender essas variações permite maior acurácia na previsão de riscos e maior efetividade nas ações policiais.

1.5.2 Modelagem Estratégica e Análise Preditiva

A próxima fase consistirá no desenvolvimento da modelagem do planejamento estratégico, voltada à previsão de crimes e à alocação racional de recursos. Serão definidos os parâmetros e variáveis relevantes à segurança pública, além da escolha das técnicas computacionais adequadas para a modelagem.

Será construída uma ontologia específica para a área de Segurança Pública, que dará suporte à representação semântica dos dados e permitirá o cruzamento inteligente de informações operacionais.

1.5.3 Desenvolvimento do Modelo Preditivo de Localização de Patrulhas

Por fim, será desenvolvido o modelo de localização preditiva das patrulhas policiais. Esse modelo utilizará dados históricos e em tempo real para identificar áreas de maior risco e otimizar a distribuição das viaturas. Serão avaliadas diferentes técnicas de agrupamento e localização, tais como os modelos RTM, k-means, p-medianas, entre outros, com vistas à construção de mapas dinâmicos de calor criminal e planos de patrulhamento mais eficazes.

1.6 Organização

A dissertação está organizada da seguinte forma, numerando-se de acordo com os capítulos:

1. **Proposta do Projeto** - Apresenta as considerações iniciais sobre segurança pública e a necessidade de aprimorar o policiamento por meio de tecnologias avançadas. São discutidos os problemas, hipóteses e objetivos da pesquisa, além da fundamentação teórica que embasa o estudo.

2. **Processo de Extração do Conhecimento e Modelo de Aprendizado** - Explora o processo de extração de conhecimento e os métodos de Aprendizado de Máquina empregados na pesquisa. São detalhadas as etapas de seleção, pré-processamento, transformação, mineração de dados e interpretação dos resultados.

3. **Policiamento Inteligente e Tecnologias Emergentes** - Introduz conceitos sobre policiamento inteligente e o impacto de tecnologias emergentes na segurança pública. São abordados modelos como Risk Terrain Modeling (RTM), K-means e policiamento preditivo, além dos desafios financeiros para implementação dessas soluções.

4. **Ontologias, Redes Neurais e Análise de Dados** - Descreve o uso de ontologias para aprimorar a representação de dados, redes neurais convolucionais (CNNs) para reconhecimento de padrões, e a análise preditiva para antecipação de eventos criminais. Discute-se a relação entre Aprendizado de Máquina e segurança pública.

5. **Predição Criminal Utilizando Técnicas de Machine Learning e Análise Espacial** - Apresenta a aplicação de técnicas de Aprendizado de Máquina para previsão criminal, com destaque para métricas de desempenho, modelagem de policiamento inteligente e análise espacial. São discutidas as funções de perda utilizadas para otimização dos modelos.

6. **Conclusão** - Resume os principais achados da pesquisa, valida a hipótese inicial e discute as respostas às perguntas formuladas. São propostas direções para trabalhos futuros e melhorias na implementação prática dos modelos estudados.

CAPÍTULO 2

Processo de Extração do Conhecimento e Modelo de Aprendizado

2.1 Considerações Iniciais

A implementação bem-sucedida de sistemas de policiamento preditivo baseados em Inteligência Artificial (IA) depende significativamente da coleta, pré processamento e análise exploratória de dados. Essas etapas desempenham um papel importante desde a obtenção inicial dos dados até a preparação para aplicação de algoritmos de aprendizado de máquina. Além disso, este capítulo explora o processo de extração de conhecimento, conforme definido por (U.M, 1996), e investiga as técnicas de policiamento preditivo no contexto do policiamento público.

O policiamento preditivo envolve técnicas analíticas que identificam "alvos prováveis para intervenção policial"(Perry, 2013). O primeiro simpósio sobre policiamento preditivo do *National Institute of Justice*, realizado em Los Angeles em 2009, identificou inúmeras aplicações potenciais, destacando especialmente o uso do tempo e local de futuras incidências em padrões ou séries de crimes. Outra definição relevante para o policiamento de rua é "o uso de dados históricos para criar uma previsão espaço-temporal de áreas de criminalidade ou pontos críticos de crime que serão a base para decisões de alocação de recursos policiais, com a expectativa de que ter policiais no local e horário propostos deterá ou detectará atividades criminosas"(Ratcliffe, 2010).

O estudo atual usa o policiamento preditivo para explorar quais atividades de patrulha geográfica devem ser perseguidas nessas áreas prioritárias. Isso

é relevante não apenas para muitos departamentos de polícia, mas também para o avanço de nossa compreensão teórica do policiamento preditivo e de um conceito estreitamente relacionado, o policiamento de pontos críticos.

Infelizmente, a pesquisa existente é menos clara sobre quais atividades específicas os policiais devem realizar nos pontos críticos. A ideia por meio do algorítimo é gerar grades de missão, cada uma com duração de 3 (três) meses. No primeiro período experimental aplicar o algoritmo e as respostas de patrulha aos crimes contra o patrimônio, e o segundo, deixar de aplicar a predição para testar a acurácia. Da mesma forma fazer de modo inverso em área que contenha características semelhantes e ao final testar a real contribuição do algorítimo.

2.2 Processo de Extração de Conhecimento

Dentro desse contexto, o processo de extração de conhecimento, conforme definido por (U.M, 1996), desempenha um papel crucial. A extração de conhecimento, ou *Knowledge Discovery in Databases (KDD)*, envolve várias etapas, incluindo a seleção, pré-processamento, transformação, mineração de dados e interpretação/avaliação dos dados. No âmbito do policiamento preditivo, a KDD é aplicada para transformar grandes volumes de dados históricos e em tempo real em informações acionáveis. Por exemplo, ao analisar padrões de crimes passados e condições atuais, o KDD pode identificar tendências ocultas e anomalias que informam as grades de missão geradas pelo algoritimo proposto. Isso não apenas aprimora a precisão das previsões, mas também fornece uma base robusta para decisões estratégicas de alocação de recursos. Ao integrar técnicas de mineração de dados com algoritmos de policiamento preditivo, é possível desenvolver modelos mais sofisticados que capturam a complexidade e dinâmica dos ambientes urbanos, facilitando intervenções policiais mais eficazes e oportunas.

Dentro do contexto da análise preditiva criminal, o processo de extração de conhecimento seria aplicado para transformar vastas quantidades de dados históricos e em tempo real em informações acionáveis que aprimoram a eficácia do policiamento preditivo. A extração de conhecimento envolveria várias etapas: primeiro, a seleção e pré-processamento dos dados de crimes, condições meteorológicas, eventos especiais e dados socioeconômicos; em seguida, a transformação desses dados para um formato adequado para análise; depois, a mineração de dados utilizando algoritmos avançados de Aprendizado de Máquina para identificar padrões ocultos e tendências; e, finalmente, a interpretação e avaliação dos resultados para gerar previsões espaço-temporais precisas sobre onde e quando crimes podem ocorrer. Essas informações são então utilizadas para criar grades de missão, orientando as patrulhas policiais

Figura 2.1: Processo de extração do conhecimento. Fonte adaptada de Fayyad(1996).

de forma mais eficiente e informada, melhorando a alocação de recursos e a tomada de decisão estratégica no combate ao crime.

2.2.1 Seleção de Dados:

A primeira fase envolve a seleção de dados, que pode incluir registros históricos de crimes, dados demográficos, condições meteorológicas, e outras informações relevantes para o contexto do policiamento. Métodos modernos de seleção podem incluir a integração automática de sistemas de registro policial, sensores urbanos e fontes de dados abertos, proporcionando uma visão abrangente e em tempo real do ambiente urbano.

2.2.2 Pré-processamento de Dados:

Após a seleção, os dados são submetidos a um rigoroso processo de pré processamento. Esta etapa inclui limpeza dos dados para remover valores ausentes ou inconsistentes, normalização para garantir que todas as variáveis estejam na mesma escala, e transformação para formatos adequados para análise. A seleção de características relevantes também ocorre nesta etapa, identificando quais variáveis serão utilizadas nos modelos preditivos. O pré-processamento é extemamente necessário para garantir a qualidade dos dados antes da análise exploratória e da aplicação de algoritmos de IA.

2.2.3 Transformação de Dados:

A transformação de dados envolve a preparação dos dados para análise detalhada. Isso pode incluir a redução da dimensionalidade, a codificação de variáveis categóricas, e outras técnicas para tornar os dados mais adequados aos métodos analíticos a serem utilizados.

2.2.4 Mineração de Dados:

A mineração de dados visa entender melhor as características e padrões nos dados selecionados. Técnicas como visualizações gráficas, estatísticas descritivas e análise de correlação são empregadas para identificar tendências, *outliers* e relações entre variáveis. Essa etapa não apenas prepara os dados para modelagem preditiva, mas também proporciona *insights* iniciais que podem orientar decisões sobre quais técnicas de Aprendizado de Máquina são mais apropriadas para o problema em questão.

2.2.5 Interpretação e Avaliação:

Ao implementar sistemas de policiamento preditivo, podemos enfrentar desafios técnicos como a integração de múltiplas fontes de dados, a garantia da qualidade dos dados e a escolha adequada de técnicas de pré-processamento e análise. Ao seguir metodologias rigorosas nessas áreas, as agências policiais podem não apenas melhorar a eficiência na prevenção e combate ao crime, mas também avançar na compreensão das dinâmicas complexas das cidades modernas. Este capítulo oferece uma base sólida para o desenvolvimento e aprimoramento contínuo de soluções de segurança pública baseadas em dados e IA.

2.3 Modelo de aprendizado

Uma das características mais notáveis do ser humano é a capacidade de aprender. Esse aprendizado ocorre por meio de ensinamentos ou das experiências vividas. O aspecto mais significativo do aprendizado humano é a habilidade de aplicar o conhecimento adquirido a problemas inéditos.

Os computadores são máquinas notáveis, com capacidades de processamento e memória que superam as dos seres humanos. No entanto, eles têm uma limitação significativa: não são capazes de tomar decisões de forma independente. Se fosse possível programar computadores para que pudessem aprender e melhorar seu desempenho ao longo do tempo, isso seria extraordinário. Isso combinaria as capacidades humanas e computacionais, criando um sistema que pudesse aprender e operar com a eficiência de um computador. A aprendizagem de máquina (AM) é uma área multidisciplinar que busca capacitar os computadores a aprenderem (Michalski et al., 1986) . Esta área é multidisciplinar porque incorpora resultados de Inteligência Artificial, estatística, probabilidade, teoria da complexidade computacional, teoria da informação, psicologia, filosofia, neurobiologia, entre outros campos.

Diz-se que um computador aprende quando o uso de experiências na resolução de um conjunto de tarefas melhora com a exposição a essas experiências, ou seja, ele pode aprimorar seu desempenho com base nas experiências

adquiridas. A aprendizagem dos computadores ocorre quando eles conseguem formar ações generalizadas a partir de experiências prévias. Essas experiências são apresentadas na forma de dados e, através de algoritmos, é possível aprender, ou seja, identificar e construir padrões gerais presentes nos dados.

Os testes para medir a capacidade de aprendizagem do computador são realizados observando-se uma nova experiência e comparando esse resultado com os obtidos em experiências anteriores.

Dentro do contexto do policiamento preditivo e da integração de IA, o processo de extração de conhecimento é essencial para o desenvolvimento e aprimoramento dos modelos de aprendizado utilizados, como os implementados no algorítimo de predição criminal. A extração de conhecimento envolve a identificação e transformação de dados brutos em informações significativas e acionáveis, um passo crítico para o sucesso dos modelos de aprendizagem de máquina. Ao aplicar técnicas de (KDD), como seleção, pré-processamento, transformação e mineração de dados, os analistas podem revelar padrões ocultos e correlações complexas nos dados de crimes históricos e contextuais. Esses *insights* são então integrados nos modelos de aprendizado, como redes neurais convolucionais e algoritmos de árvores de decisão, entre outros, que são treinados para reconhecer e prever padrões futuros de atividade criminosa.

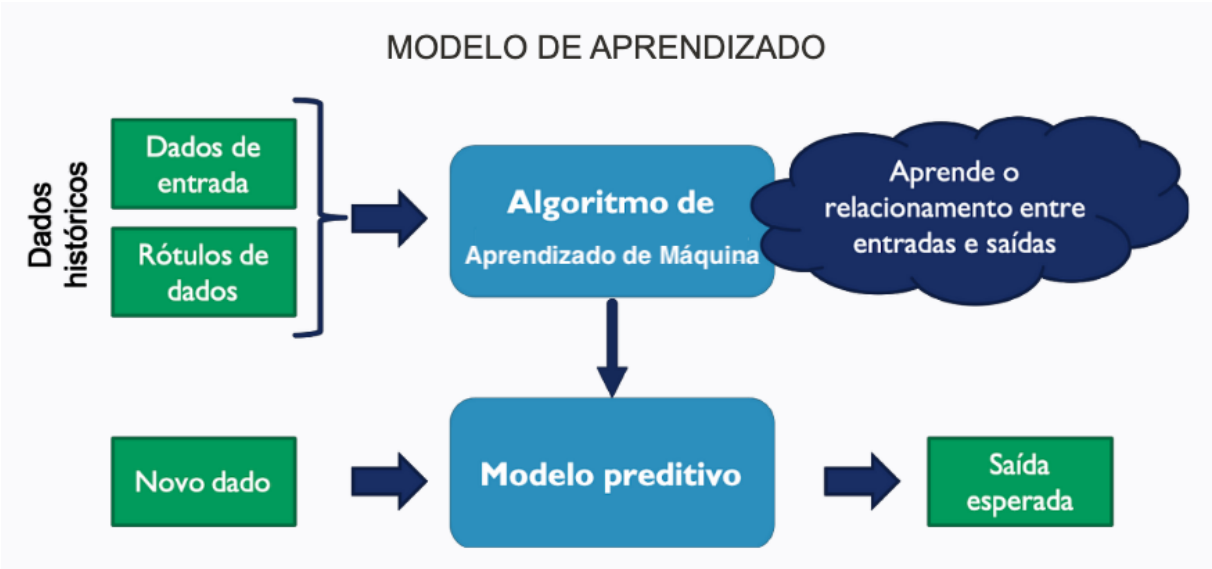


Figura 2.2: Modelo de Aprendizado. Fonte adaptada de Escovedo(2020).

A figura 2.2 apresenta um fluxo de trabalho típico no contexto do policiamento preditivo utilizando aprendizado de máquina. Inicialmente, os dados históricos são coletados e preparados como entrada para o algoritmo de machine learning. Esses dados incluem informações sobre crimes passados, como tipos de crimes, localização, horário e outras variáveis contextuais rele-

vantes, que são rotuladas para indicar a ocorrência ou não de crime em um determinado período. O algoritmo de machine learning, como uma rede neural convolucional ou um modelo de árvore de decisão, é então treinado com esses dados históricos para aprender padrões complexos e relações entre variáveis.

À medida que novos dados são coletados, como atualizações de registros criminais ou mudanças nas condições climáticas, urbanas, trânsito, eventos, períodos sazonais de feriados, finais de semana prolongados, época de pagamentos, etc, esses dados são processados pelo modelo preditivo. O modelo aplica os padrões aprendidos durante o treinamento para gerar previsões sobre onde e quando novos crimes podem ocorrer. A saída esperada do modelo preditivo é uma análise espacial e temporal que identifica áreas de alto risco de criminalidade, representadas visualmente em mapas interativos ou relatórios detalhados. Essas previsões serão utilizadas pelos gestores de policiamento para alocar recursos de maneira estratégica, otimizando a presença policial e melhorando a eficiência na resposta a incidentes criminais.

2.4 Impactos e Considerações finais

Entendemos que integração de Inteligência Artificial, extração de conhecimento e policiamento preditivo representará um grande avanço dentro do contexto da ciência da computação e especialmente no campo do policiamento público. A aplicação de técnicas avançadas de aprendizado de máquina, como exemplificado pelo predição criminal, tende a capacitar as forças policiais a reagirem de forma mais eficaz a crimes, além de prever e prevenir incidentes antes que ocorram. Utilizando dados históricos e algoritmos acertivos, esses sistemas melhoram não só a alocação de recursos, mas também a capacidade de resposta das forças policiais diante das dinâmicas complexas das cidades modernas.

No entanto, é importante reconhecer as limitações tecnológicas que podem surgir. A precisão dos modelos de policiamento preditivo depende da qualidade e da disponibilidade e qualidade dos dados utilizados, bem como da capacidade dos algoritmos em lidar com variações e mudanças no ambiente urbano. O desenvolvimento futuro dessas tecnologias deve concentrar-se não apenas na inovação técnica, mas também em superar essas limitações para garantir a eficácia contínua do policiamento preditivo.

CAPÍTULO

3

Policiamento Inteligente e tecnologias emergentes

3.1 Considerações Iniciais

Neste capítulo são apresentados o contexto e a motivação para a proposta deste projeto de mestrado: segurança pública e o apoio da Tecnologia da Informação (TI) à sua prática. Por se tratar de uma proposta de projeto que utilizará dados da Polícia Militar do Estado de São Paulo, é apresentada uma breve descrição da situação atual quanto à política de uso de TI na Segurança Pública, e sua base legal.

3.2 Policiamento Inteligente

No cenário em constante evolução da segurança pública, surge uma abordagem inovadora que visa a aprimorar as práticas tradicionais de policiamento: o policiamento inteligente. Diante da necessidade de enfrentar os desafios contemporâneos relacionados ao crime e à preservação da ordem pública, é imperativo explorar novas estratégias que incorporem tecnologias emergentes para otimizar a eficiência das forças de segurança.

O policiamento ostensivo, marcado pela presença visível das autoridades, sempre foi uma resposta da sociedade ao crime, refletindo a autorização coletiva para o uso legítimo da força quando as normas de convivência social são transgredidas. No entanto, a busca por aprimoramento dessas práticas conduz ao advento do policiamento inteligente, que se baseia na integração de tecnologias avançadas para prevenção, resposta e controle mais eficazes.

No âmbito da PMESP, as Normas para o Sistema Operacional de Policiamento PM (NORSOP) definem as ações de polícia ostensiva e de polícia de

Figura 3.1: Plano de Policiamento. Fonte Apresentação Fiesp (Cmt Geral 2019)

preservação da ordem pública, conforme observa-se abaixo:

> "6.3.1.1. polícia ostensiva conceito abrangente, que envolve atividades de prevenção primária e secundária, as quais são executadas para consecução da segurança pública, tais como policiamento comunitário, radiopatrulhamento e todas as demais que são levadas a efeito pela Polícia Militar a fim de prevenir o cometimento de ilícitos penais ou de infrações administrativas sujeitas ao controle da Instituição.
> 6.3.1.2. polícia de preservação da ordem pública é a atividade cometida à Polícia Militar de restauração da ordem pública, envolvendo a repressão imediata às infrações penais e administrativas e a aplicação da lei (PMESP2022, 2022)."

O Comando da Instituição, por intermédio do Plano de Comando 2022/ 2023 (PMESP2022, 2022), estabeleceu como visão de futuro da Instituição

ser reconhecida no cenário nacional e internacional como uma referência na prestação de serviços de segurança pública.

No mesmo sentido, o Comando da Instituição, por intermédio da 6ł Seção do EM/PM (6ł EM/PM), elaborou o Planejamento Estratégico da Instituição para um período de 20 anos.

A 6ł EM/PM define Planejamento Estratégico como: conjunto de atos ordenados destinados a‘ validação dos princípios organizacionais, construção de cenários, identificação de objetivos estratégicos, indicadores e metas.

Nesse cenário de consolidação da Planejamento e visão de Futuro da PMESP, torna-se oportuno abordar alguns desafios estratégicos vislumbrados, conforme segue:

1. gestão com foco nos *Partes Interessadas*, geração de valor e diferenciação;
2. gestão da demanda pela chamada emergencial 190;
3. gestão da carência de Recursos Humanos;
4. gestão da restrição de recursos orçamentário;
5. gestão da Tecnologia da Informação e Comunicação.

As inovações que se processaram no âmbito da PMESP ao longo do tempo, muitas vezes à frente da evolução natural do ambiente social, geraram na Instituição uma percepção sobre a necessidade de criar e manter um harmônico e eficiente funcionamento de seus sistemas operacionais e administrativos. Assim, para atingir seus objetivos globais com eficiência, a PMESP elaborou o Sistema de Gestão da Polícia Militar do Estado de São Paulo (GESPOL), com o objetivo de atender às necessidades e expectativas de seus públicos de relacionamento, bem como cumprir com excelência os serviços que realiza, com base no pensamento sistêmico e no suporte doutrinário.

Segundo o GESPOL (PMESP2022, 2022), o suporte doutrinário está alicerçado sobre os princípios de Polícia Comunitária, de Direitos Humanos e de Gestão pela Qualidade Ao se debruçar sobre as Teorias de Administração Pública, alguns autores identificam a presença de uma nova administração pública voltada para a eficiência, a eficácia e a efetividade do aparelho do Estado com foco em resultados, fruto da percepção, adaptação e, até mesmo, antecipação as mudanças potenciais dos administradores públicos. Diante do exposto, não resta dúvida de que as Instituições públicas ou privadas na sociedade moderna têm consciência da importância dos resultados obtidos com a gestão com foco nos *Partes Interessadas*, geração de valor e diferenciação, pois no mundo moderno não basta ter capital, tecnologia, recursos materiais e humanos, é preciso desenvolver processos que possibilitem ser eficaz e eficiente.

As tecnologias emergentes desempenham um papel fundamental nesse novo paradigma de policiamento. Na vanguarda do avanço tecnológico, o policiamento inteligente abraça o potencial transformador das redes neurais convolucionais (CNNs), ontologias e análise preditiva, promovendo uma abordagem inovadora e eficiente para a segurança pública.

Essa integração de tecnologias avançadas não apenas reforça a capacidade das forças policiais, mas também redefine o panorama do policiamento, permitindo um planejamento estratégico mais preciso e uma roteirização eficiente das operações.

As redes neurais convolucionais, inspiradas na estrutura do cérebro humano, emergem como ferramentas poderosas no processamento e interpretação de dados visuais. No contexto do policiamento, essas redes podem ser aplicadas na análise de imagens de vigilância, reconhecimento facial e identificação de padrões comportamentais. A capacidade de identificar de forma rápida e precisa elementos relevantes em grandes conjuntos de dados visuais amplifica a eficácia do monitoramento urbano, permitindo uma resposta mais ágil a atividades suspeitas.

A utilização de ontologias, que representam formalmente o conhecimento em domínios específicos, oferece uma estrutura semântica robusta para a compreensão e organização dos dados relacionados à segurança pública. Ao criar uma base de conhecimento compartilhada, as ontologias facilitam a interoperabilidade entre diferentes sistemas e fontes de informação. Isso possibilita uma compreensão mais holística do ambiente de segurança, auxiliando na identificação de relações e padrões complexos que podem passar despercebidos em abordagens convencionais.

A ideia de trabalho com análise preditiva, aliada às redes neurais e à ontologia, pressupõe a utilização de algoritmos avançados que processam dados históricos e em tempo real, identificando padrões que indicam áreas de maior probabilidade de ocorrência de crimes. Essa capacidade preditiva busca antecipação de ameaças e também o planejamento estratégico, a alocação de recursos e a roteirização eficiente das operações policiais.

Além disso, a conectividade e a comunicação aprimoradas entre as agências de segurança promovem uma resposta mais rápida e coordenada a situações de emergência. A implementação de sistemas integrados permite a troca instantânea de informações entre diferentes setores, fortalecendo a capacidade de resposta diante de eventos críticos.

Entre as ações para a garantia da segurança pública, a atividade de policiamento, através da presença ostensiva apresenta-se como o aspecto mais visível da reação da sociedade ao crime e a preservação da ordem pública, sendo uma a medida de controle social necessária para o bem-estar e con-

vivência harmônica (Cordner et al., 2007). Esse policiamento tem entre suas principais características a autorização coletiva social para o uso real da força, pelos agentes públicos que a exercem, quando alguém transgride as normas de convivência social (Bayley, 2002). Logo, os modelos tradicionais de policiamento buscam, através de ações preventivas e reativas, evitar que o crime ocorra ou reestabelecer o direito violado (integridade física e patrimônio), diante de uma agressão real ou iminente.

De tal modo, busca-se que os criminosos, a partir da presença policial, não cometam ilícitos (Koper, 1995), sendo fundamental a alocação eficiente dos recursos existentes, para que os delitos possam ser evitados, através da presença policial no local propício ao acontecimento do fato criminal (Braga et al., 2000, Basilio et al., 2019).

Com esse objetivo, se faz necessário que os recursos sejam empregados nos locais onde possam evitar os fatos infracionais ou atender de forma mais rápida aos acontecimentos de emergência, sendo que a busca por soluções de otimização da alocação dos recursos policiais, seja na disposição de postos policiais ou viaturas, seja na definição de roteirização, têm sido objeto de estudo por décadas (Simpson e Hancock, 2009, Dewinter et al., 2020), a partir do desenvolvimento e aplicação de problemas clássicos como o Problema de Localização de Facilidades e Problema de Roteamento de Veículos.

3.3 Tecnologias emergentes

Os modelos tradicionais de policiamento buscam, através de ações preventivas e reativas, evitar que o crime ocorra ou reestabelecer o direito violado (integridade física e patrimônio), diante de uma agressão real ou iminente.

Nesse sentido, diversos métodos de policiamento têm sido desenvolvidos, com fito de identificar, mormente, onde e quando ocorrerão os eventos criminais, para que a partir da alocação dos recursos limitados, nos locais de maior risco, esses crimes previstos possam ser evitados (Perry, 2013).

3.3.1 Risk Terrain Modeling (RTM)

A análise de risco proposta, introduzida inicialmente no Brasil no Estado do Rio Grande do Sul e posteriormente no Estado de São Paulo na cidade de Barueri, envolve a compilação, em um mapa composto, de camadas que representam diferentes aspectos de influência criminal, permitindo avaliar tanto a influência quanto a intensidade desses fatores na ocorrência de determinados crimes. Dessa forma, essa abordagem atua como uma ferramenta analítica que auxilia na identificação de atributos do ambiente que estão espacialmente associados à ocorrência de crimes (Valasik et al. (2021)).

Em uma análise preliminar, os resultados obtidos podem ser qualitativamente semelhantes aos gerados por métodos de hotspot, os quais destacam

áreas com elevado índice de criminalidade. No entanto, do ponto de vista analítico, essas metodologias são distintas; enquanto as análises de hotspot utilizam técnicas de agrupamento que refletem eventos criminais passados, o Risk Terrain Modeling (RTM) adota uma abordagem de classificação, caracterizando o risco de criminalidade de uma região com base em suas características geográficas (Perry (2013)).

O diferencial do modelo reside na correlação entre o "risco de crime"e a "oportunidade para o crime", através da avaliação dos graus de risco de diferentes locais, levando em conta fatores criminógenos em comparação com outros pontos dentro da mesma unidade espacial padronizada.

O processo do Risk Terrain Modeling (RTM) se estrutura em três elementos principais: "1) padronização dos conjuntos de dados em uma geografia comum, 2) diagnóstico dos fatores de risco no espaço e 3) articulação das vulnerabilidades espaciais"(Kennedy et al. (2011)). Essa abordagem permite uma análise mais precisa e contextualizada, contribuindo para a compreensão das dinâmicas criminais em áreas específicas e auxiliando na elaboração de estratégias de prevenção mais eficazes.

Nesse ponto, inicia-se a abordagem técnica do Modelo de Análise de Risco Territorial, com a identificação preliminar dos fatores associados a um crime específico cujo risco está sendo avaliado. Essa identificação pode ser realizada por meio de meta-análise, métodos empíricos, revisão de literatura, e pela experiência e conhecimento do profissional envolvido (Caplan et al. (2011)).

Após a seleção dos fatores relevantes, procede-se à plotagem padronizada dos dados em um ambiente geográfico comum. Essa etapa envolve a criação de uma camada no mapa que representa a área definida, com as células de análise delimitadas. Cada fator identificado é, então, representado por um mapa de cobertura (raster) que utiliza a mesma geografia (Caplan et al. (2011)).

Com isso, ocorre a combinação das camadas de risco, resultando em um mapa composto, ou mapa de risco do terreno. Neste mapa, a cada local (célula/grid) na área definida é atribuído um valor de risco, que considera todos os fatores relacionados ao crime específico. Essa representação visual permite uma análise mais intuitiva e informada dos riscos associados, facilitando a tomada de decisões em políticas de segurança pública.

O mapa de risco combinado identifica, assim, as células/grids que apresentam o maior potencial de ocorrência do crime em questão. Nesse contexto, ele funciona como um guia para o planejamento tático-operacional, auxiliando na alocação de recursos policiais e na formulação de estratégias e ações que visem a diminuição do risco local. Essa redução pode ser obtida por meio da dissuasão proporcionada pela presença policial ou por intervenções estruturais que transformem o ambiente de risco.

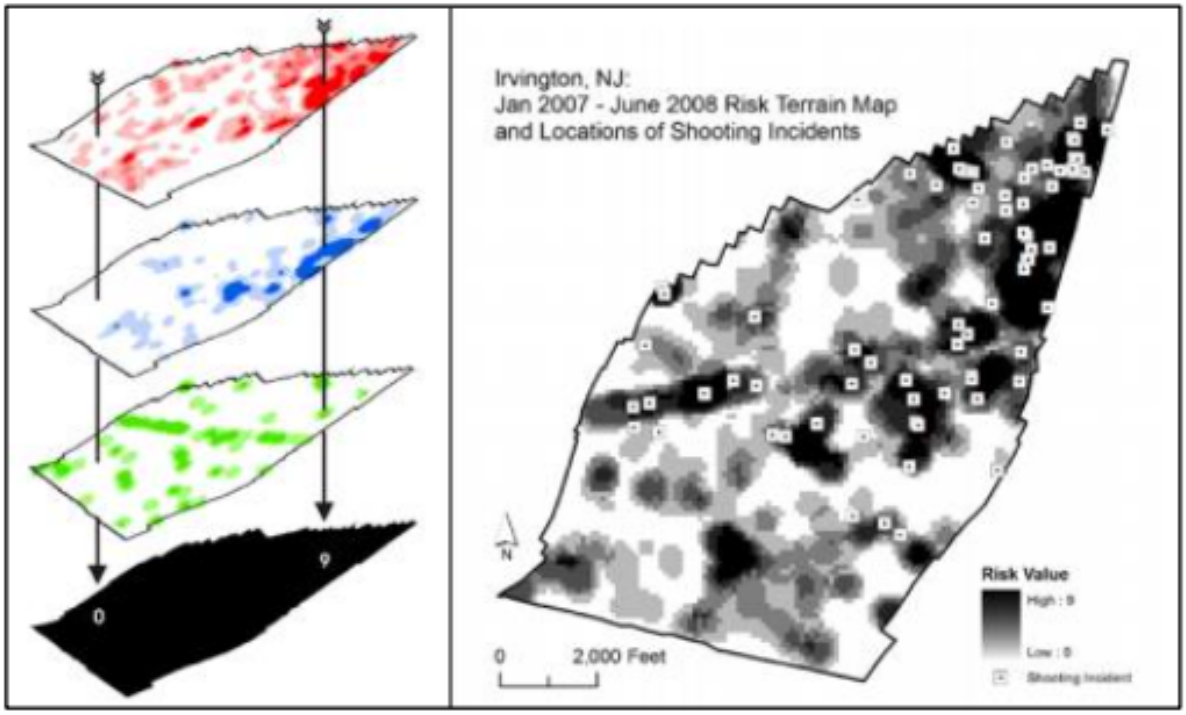


Figura 3.2: Representação de combinação de camadas de risco - Fonte: Caplan et al. (2011)

Dessa maneira, a análise preditiva do Risk Terrain Modeling (RTM) se torna uma ferramenta valiosa para a adoção de estratégias dentro do policiamento preditivo, alinhando-se aos conceitos discutidos. Essa abordagem não apenas busca uma resposta imediata ao crime, mas também visa a prevenção, mitigando as condições que favorecem sua incidência.

3.3.2 K-means)

O método de clusterização k-means é um algoritmo de Aprendizado de Máquina não supervisionado, o que significa que não requer classificação ou parametrização prévias para realizar a clusterização (Sinclair e Das (2021)). Como técnica de agrupamento, ele visa organizar dados em k classes, utilizando uma medida de similaridade para agrupar elementos que compartilham grande semelhança e identificar padrões (Jain (2010)). Para isso, o algoritmo busca determinar um conjunto de k pontos que funcionam como centros (centróides) e, em seguida, atribui cada elemento aos seus centróides mais próximos, formando os clusters (Gusmão et al. (2020)).

Segundo Jain (2010), o algoritmo k-means tem uma história rica e diversificada, pois foi descoberto independentemente em diferentes campos científicos por Steinhaus (1956), Lloyd (proposto em 1957, publicado em 1982), Ball e Hall (1965) e MacQueen (1967). Ele destaca que, apesar do longo intervalo desde sua introdução, o k-means continua sendo um dos algoritmos mais uti-

lizados para a clusterização de dados, devido à sua simplicidade, eficiência e sucesso empírico (Jain (2010); Oliveira et al. (2015); Gusmão et al. (2020); Sinclair e Das (2021)).

É importante ressaltar que o algoritmo k-means necessita da definição de três parâmetros pelo usuário: o número de clusters k, a inicialização dos clusters e a métrica de distância (Jain (2010)). No que diz respeito aos pontos de inicialização, vale a pena observar que diferentes abordagens podem levar a agrupamentos finais imprecisos, já que o k-means tende a convergir para um mínimo local.

De maneira geral, o objetivo do k-means é encontrar os clusters de forma a minimizar a soma das distâncias entre os pontos dos dados e o centróide mais próximo (Oliveira et al. (2015)). Assim, a função objetivo, de forma simplificada, pode ser expressa conforme apresentado por Gusmão et al. (2020):

$$f(x) = Min \sum_{k=1}^{k} \sum_{x_i \in C_k} d(x_i, x_{0k})$$

Figura 3.3: Função K-means - Fonte: Gusmão et al. (2020)

Onde: - x_k^0: é o centróide do cluster x_C^k; e - $d(x_i, x_k^0)$: é a distância entre o ponto x_i e o centróide x_k^0.

É importante destacar que a minimização dessa função objetivo é classificada como um problema NP-hard Jain (2010). No entanto, o k-means adota um algoritmo que apresenta boa escalabilidade, permitindo agrupar conjuntos de dados com dimensões moderadas a altas, tudo isso com custos computacionais relativamente baixos Sinclair e Das (2021).

O funcionamento do algoritmo é dividido em três etapas Allen-Zhu et al. (2019): 1. Na primeira etapa, k centroides são definidos ou selecionados aleatoriamente dentro do conjunto de dados original. 2. Em seguida, calcula-se a distância entre cada ponto da base de dados e cada centróide, vinculando cada ponto ao centróide mais próximo. O k-means normalmente utiliza a métrica euclidiana para determinar as distâncias entre pontos e centros de cluster. 3. Por fim, os centroides são atualizados para o centro dos clusters formados pelos pontos associados ao mesmo centróide. As etapas 2 e 3 são repetidas até que os centroides não mudem, encerrando assim o processo.

De acordo com Gusmão et al. (2020), esse método pode ser eficaz em problemas de localização de facilidades, já que consiste na identificação dos k centroides, o que define potenciais locais candidatos para seu posicionamento.

Adicionalmente, Allen-Zhu et al. (2019) aplicam o k-means na análise criminal, propondo um modelo preditivo que utiliza características dos crimes para prever a quantidade de ocorrências em determinados micro locais.

Por outro lado, Sinclair e Das (2021) empregaram o algoritmo de clusterização para detectar padrões de acidentes rodoviários no Reino Unido, buscando entender as relações entre as variáveis registradas pelo departamento de polícia.

3.3.3 Policiamento Preditivo (Predictive Policing)

A possibilidade de prever ou predizer os crimes parte do reconhecimento de que os crimes não ocorrem de forma homogênea no espaço geográfico, mas, encontram-se concentrados em determinadas áreas e locais, conhecidas como hotspots, por razões que podem ser explicadas em relação à interação entre vítima e infrator, e as oportunidades que existem para cometer crimes (Chainey et al., 2008, Ratcliffe, 2010)

Logo, esses pontos críticos, tradicionalmente definidos como as regiões com taxas de crime mais altas (Coldren Jr et al., 2013) em alguns casos possuindo pontos com ainda mais risco dentro desses hotspots, podem ser utilizados para decifrar os padrões espaciais do crime.

A partir da percepção e identificação dos locais de concentração de crimes, estudos apontam que a realização de intervenções policiais focadas aos pontos críticos, como patrulhas dirigidas, prisões proativas e policiamento orientado para o problema, concentrando os recursos nas microunidades de maior risco (Hart e Zandbergen, 2014), pode produzir ganhos significativos na prevenção do crime, a partir de dois mecanismos teóricos fundamentais de prevenção do crime: dissuasão e redução das oportunidades de crime (Braga et al., 2000).

Entre os trabalhos sobre a eficácia das ações policiais (Koper, 1995), apresentou um estudo amplamente citado, no qual destaca que o patrulhamento intermitente com período de paradas em pontos de altos índices de criminalidade (segmentos de rua ou quarteirões). Segundo o estudo, analisando ações de polícia nas modalidades de patrulhamento (deslocamento da viatura) e permanência (paradas em pontos estratégicos). Ainda, a partir dos dados demonstra que após um período de parada, mesmo após a saída da viatura permanece um período de manutenção da ordem no local, com uma probabilidade reduzida de incidentes, por algum período.

o Policiamento Preditivo pode ser definido como a aplicação de técnicas analíticas - principalmente técnicas quantitativas - para identificar alvos prováveis para intervenção policial e prevenir crimes ou resolver crimes passados através de previsões estatísticas (Perry, 2013).

A partir das definições pode-se sintetizar que policiamento preditivo é um método do processo de tomada de decisão dos gestores de segurança pública,

utilizando técnicas estatísticas e preditivas no planejamento das ações policiais proativas nos locais de maior risco, para alcançar maior eficiência (melhor emprego dos recursos), eficácia (atingindo os resultados esperados) e efetividade (gerando impacto no local de atuação).

Perry (Perry, 2013), propõem um modelo de processo do policiamento preditivo, a partir de um sistema com quatro fases (Figura 3.4): coleta de dados, análise, operações policiais e resposta criminal.

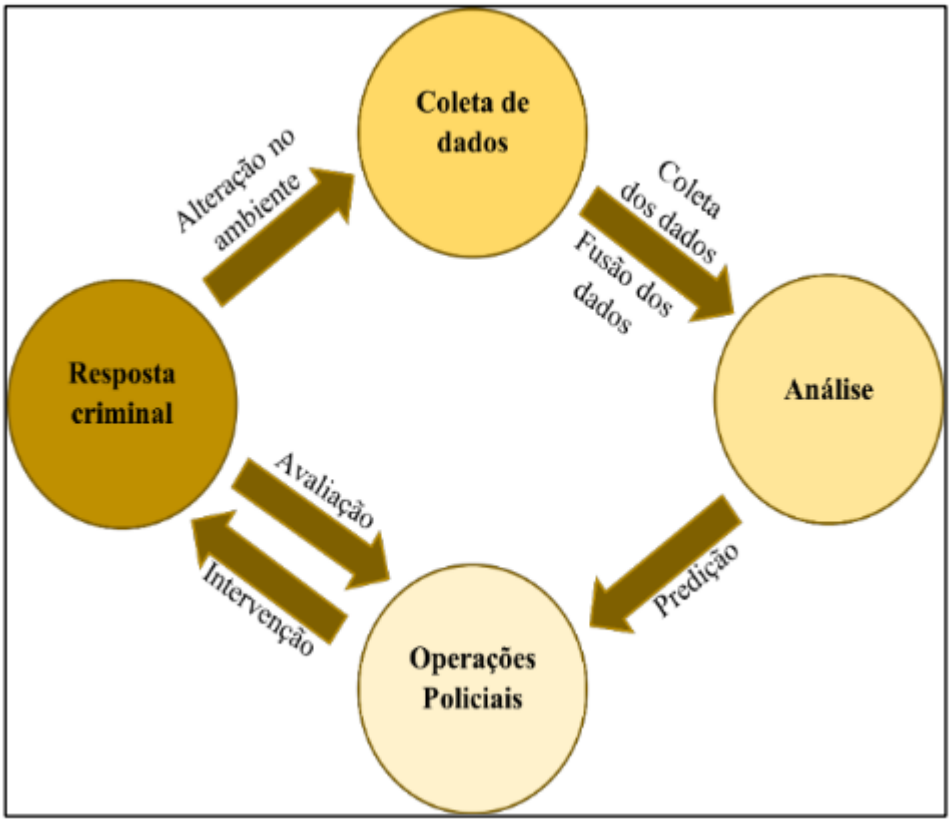


Figura 3.4: Processo do Policiamento Preditivo - Fonte Perry et al 2013

A etapa de coleta de dados é fundamental, uma vez que as técnicas de predição dependem de informações. Os bancos de dados da polícia, normalmente, contêm uma grande quantidade de dados criminais que podem ser usados identificar sobre as tendências e padrões atuais e futuros do crime (Rummens et al., 2017).

Na fase de análise, os métodos e as técnicas são implementados não apenas para identificar os fatos ocorridos e o cenário existente, mas, para antecipar os próximos crimes. Portanto, a análise constitui a base para o emprego de recursos, sendo que, se os gestores e policiais não entendem os fatores que conduzem a um aumento da chance de crime, a eficácia de suas ações pode ser reduzida ((Perry, 2013, Rummens et al., 2017).

Seguindo, é preciso que o conhecimento gerado pela análise seja considerado no planejamento e execução das operações e ações policiais, para que

de fato possam atingir os resultados, evitando o cometimento de crimes e proporcionando maior segurança à população envolvida com os pontos e situações de risco. Logo, a terceira fase do processo, é justamente a execução de operações/ações policiais, a partir da análise preditiva. Destaca-se que o policiamento preditivo apresenta grande conexão com o policiamento preventivo, sendo que os gestores da segurança pública devem trabalhar com vários atores na sociedade para reduzir e eliminar as oportunidades e os fatores que causam o comportamento criminoso (Meijer e Wessels, 2019).

A partir da intervenção policial, haverá a fase de observar a resposta criminal, a fim de identificar possível redução, deslocamento ou encerramento de crimes na área de atuação. Diante da resposta, haverá a avaliação, com ajustes necessários nas intervenções, de forma sistemática, até que ocorra a alteração no ambiente de atuação. Isso ensejará a coleta de dados, para nova análise e planejamento, formando um ciclo da gestão preditiva do policiamento (Perry, 2013).

As abordagens preditivas são diversas, conforme cada realidade e tipos de crimes, a partir de fundamentos teóricos e técnicas estatísticas complexas. Os primeiras estudos e aplicações partiram da identificação e mapeamento de pontos críticos de crime (hotspots), a partir de mapas de densidade com base na análise retrospectiva dos dados do crime, com o consequente direcionamento dos recursos policiais para esses locais, sendo reconhecida como uma técnica eficaz de combate ao crime (Braga et al., 2000).

Do exposto, percebe-se que uma modelagem de policiamento inteligente, portanto, deve considerar a análise para identificação das características criminais no ambiente e no tempo, para antecipar o seu acontecimento, mas depende, fundamentalmente, do planejamento de ações policiais com base nessas informações, considerando a otimização dos recursos. Dessa forma, os métodos e modelos de predição devem agregar às pesquisas a identificação de pontos de maior risco e definição da forma de cobertura, através da maior permanência nesses locais.

A seguir, serão apresentados aspectos quanto aos desafios do financiamento desse tipo técnica a ser empregada no Setor Público.

3.3.4 Financiamento das tecnologias emergentes no contexto da Segurança Pública: um desafio a ser superado

Nesse contexto há que se entender o desafio financeiro, nesse sentido A Lei de Diretrizes Orçamentárias (LDO) estabelece as metas e prioridades da administração pública estadual e dispõe sobre critérios e normas que garantam o equilíbrio das receitas e despesas do Orçamento do Estado. Com base na LDO aprovada pelo Legislativo, o Poder Executivo elabora a proposta orçamentária

para o ano seguinte, em conjunto com as secretarias, autarquias e fundações, fundos especiais e as unidades orçamentárias dos poderes Legislativo, Judiciário e Ministério Público, como também das empresas estatais.

No âmbito do desafio financeiro que permeia a gestão pública, é essencial compreender o papel crucial desempenhado pela Lei de Diretrizes Orçamentárias (LDO) na definição de metas e prioridades para a administração pública estadual. A LDO não apenas delineia as diretrizes estratégicas, mas também estabelece critérios e normas que visam assegurar o equilíbrio entre as receitas e despesas do Orçamento do Estado.

Já a Lei nž 4.320, de 17 de março de 1964, e alterações posteriores, que estatui as normas gerais de direito financeiro para elaboração e controle dos orçamentos e balanços da União, dos Estados, dos Municípios e do Distrito Federal, estabelece a Lei de Orçamento como instrumento para a discriminação da receita e despesa, conforme observa-se no teor de seu artigo 2ž, reproduzido abaixo:

> "Art. 2ř A Lei do Orçamento conterá a discriminação da receita e despesa de forma a evidenciar a política econômica financeira e o programa de trabalho do Governo, obedecidos os princípios de unidade universalidade e anualidade."(LEI4320, 1964)

Nesse diapasão, a Lei Orçamentária Anual (LOA)(LOA, 2023) orça a receita e fixa a despesa do Estado para o exercício anual, compreendendo, nos termos do artigo 174, ğ 4ž, da Constituição Estadual.(CONSTITUICAOSP, 1989)

> "Artigo 174 - Leis de iniciativa do Poder Executivo estabelecerão, com observância dos preceitos correspondentes da Constituição Federal: I - o plano plurianual; II - as diretrizes orçamentárias; III - os orçamentos anuais. [...] Parágrafo 4 - A lei orçamentária anual compreenderá: 1 - o orçamento fiscal referente aos Poderes do Estado, seus fundos, órgãos e entidades da administração direta e indireta, inclusive fundações instituídas ou mantidas pelo Poder Público; 2 - o orçamento de investimentos das empresas em que o Estado, direta ou indiretamente, detenha a maioria do capital social com direito a voto; 3 - o orçamento de seguridade social, abrangendo todas as entidades e órgãos a ela vinculados, da administração direta e indireta, bem como os fundos e fundações instituídas ou mantidas pelo Poder Público;"(CONSTITUICAOSP, 1989)

Assim, todas as ações do Estado são disciplinadas pela LOA, de forma que nenhuma despesa pública pode ser executada fora do planejado, nem tão pouco fora das prioridades contidas no Plano Plurianual (PPA). Cumprindo um rito anual, em observância à CE, o Chefe do Executivo estadual encaminha o projeto ao Legislativo para análise da Comissão de Finanças, Orçamento e Planejamento (CFOP). Depois de aprovado, o Projeto da Lei Orçamentária Anual (LOA), a qual reúne todas as receitas e despesas que o governo do Estado pretende realizar no ano vindouro é sancionado pelo governador e se transforma em lei.

Esse cenário indica uma tendência de escassez de investimento de recursos por parte do Poder público na Ação Inteligência Policial Fonte Tesouro, cenário esse que sinaliza aos gestores a importância de traçar estratégias de curto e médio prazo para a captação de recursos financeiros de outras fontes e através da consolidação de parcerias, quer entre outras esferas, quer junto a iniciativa privada. Ficando claro portanto que para a PMESP obter sucesso e qualidade no cumprimento de suas missões: i) proteger as pessoas; ii) fazer cumprir a lei; iii) combater o crime; e iv) preservar a ordem pública, há a necessidade de realizar a gestão dos sistemas e infraestrutura de Tecnologia da Informação e Comunicação (TIC).

3.4 Considerações Finais

Esse cenário de contração do orçamento para investimento em TIC, obriga a PMESP a adotar processos de boas práticas de gestão e de planejamento. Entre inúmeras definições e abordagens sobre o planejamento, merecem atenção os entendimentos trazidos pelo Guia de PDTIC do Sistema de Administração dos Recursos de Tecnologia da Informação (SISP) - versão 2.0:

> "Em resumo, planejar significa orientar ações presentes e futuras, visando atingir um objetivo. O planejamento provê condições de maior segurança e menor margem de erros. É o planejamento que define ações, projetos, procedimentos, metas e objetivos, visando mudar uma situação atual ou explorar uma possibilidade futura. (pme, 2018a) [...] Permite focalizar os esforços onde os benefícios são maiores ou onde há maior necessidade (eficácia e efetividade), aproveitar melhor os recursos disponíveis, minimizando o desperdício (eficiência e economicidade), aumentar a inteligência organizacional por meio de aprendizado e responder mais adequadamente às mudanças do ambiente."(pme, 2018b)

Por fim e diante do acima exposto, as boas práticas de gestão e de planejamento de TIC (Técnologia da Informação e Comunicação), seja qual for o nível, são importantes ferramentas para a tomada de decisões pelos gestores, pois faz com que estejam aptos a agir com iniciativa e se adaptarem rapidamente às alterações do ambiente em que atuam.

CAPÍTULO

4

Ontologias, Redes Neurais (Artificiais) e Análise de Dados

4.1 Considerações Iniciais

Neste capítulo é apresentado o contexto das tecnologias emergentes que podem subsidiar de certa forma a aplicação de um planejamento de policiamento inteligente.

As Redes Neurais com enfoque nas Convolucionais que, inspiradas na complexidade do cérebro humano, despontam como ferramentas de processamento de informações visuais com potencial extraordinário. A capacidade dessas redes para identificar padrões complexos em imagens e dados visuais não apenas revoluciona a análise de informações, mas também inaugura novos horizontes na compreensão de contextos visuais, como os encontrados em vigilância urbana e reconhecimento de padrões comportamentais.

A utilização de Ontologia, por sua vez, oferece um arcabouço conceitual sólido, permitindo a representação formal do conhecimento em domínios específicos. Ao criar uma estrutura semântica, a ontologia facilita a compreensão aprofundada dos dados, promovendo a interoperabilidade entre sistemas e possibilitando uma visão mais holística das informações. Dessa forma, a ontologia torna-se um alicerce essencial para a integração eficaz de diversas fontes de dados, criando um ambiente propício para a análise e interpretação de informações heterogêneas, que hoje acaba sendo uma realidade na manipulação dos dados advindos da Segurança Pública.

Dessa combinação é possível gerar com qualidade a Análise Preditiva, que, terá por objetivo processar dados históricos e em tempo real, proporcionando não apenas a capacidade de reação a eventos iminentes, mas também a habi-

lidade única de antecipar cenários e comportamentos, possibilitando, assim, um planejamento estratégico mais eficiente e uma roteirização otimizada das operações.

Ao longo deste capítulo, vamos tentar entender conceitos básicos dessas tecnologias, compreendendo como a interconexão das Redes Neurais Convolucionais, Ontologia e Análise Preditiva que tende a desenhar um novo paradigma no cenário do policiamento inteligente.

4.2 Ontologia

A origem do termo "ontologia"remonta ao grego, onde "onto"está ligado a "ser"e "logia"está relacionado ao discurso escrito ou falado (Uschold e Gruninger, 1996, Gruber, 1993). Inicialmente utilizado no contexto filosófico, o conceito foi posteriormente adaptado no campo da ciência da computação, principalmente em estudos relacionados à aquisição de conhecimento e raciocínio. Gruber (1993) definiu ontologia como "uma especificação da conceituação", representando formalmente um conjunto de conceitos e relações que existem para um agente ou uma comunidade de agentes. Conforme a definição de Uschold e Grninger (Uschold e Gruninger, 1996, Gruber, 1993) (1996, p. 21), ontologia pode ser compreendida como:

> ...entendimento compartilhado de algum domínio de interesse que pode ser usado como uma estrutura unificadora para resolver os problemas acima descritos da maneira acima especificada. Uma ontologia necessariamente envolve ou incorpora algum tipo de visão de mundo em relação a um determinado domínio. Essa visão de mundo é frequentemente concebida como um conjunto de conceitos (por exemplo, entidades, atributos, processos), suas definições e suas inter-relações; isso é referido como uma conceitualização...(Uschold e Gruninger, 1996, Gruber, 1993)

A Ontologia, termo inicialmente filosófico, encontra aplicações em diversas áreas tecnológicas, incluindo Web semântica, Engenharia de Software e Arquitetura da Informação. Existem várias classificações de ontologias, sendo

a que se relaciona mais diretamente com a semântica, e que se pretende empregar nesta tese, aquela que utiliza conceitualização como critério principal (Gruber, 1993). As ontologias são categorizadas em quatro tipos: de alto nível, de domínio, de tarefa e de aplicação.

A descrição dos componentes da Ontologia abrange Indivíduos, Classes, Atributos e Relacionamentos. Nesse contexto, as classes representam generalizações de coleções, conceitos, tipos de objetos ou espécies de coisas, enquanto os indivíduos são suas realizações concretas. Entre os indivíduos, há relacionamentos específicos, e cada um possui um conjunto de atributos.

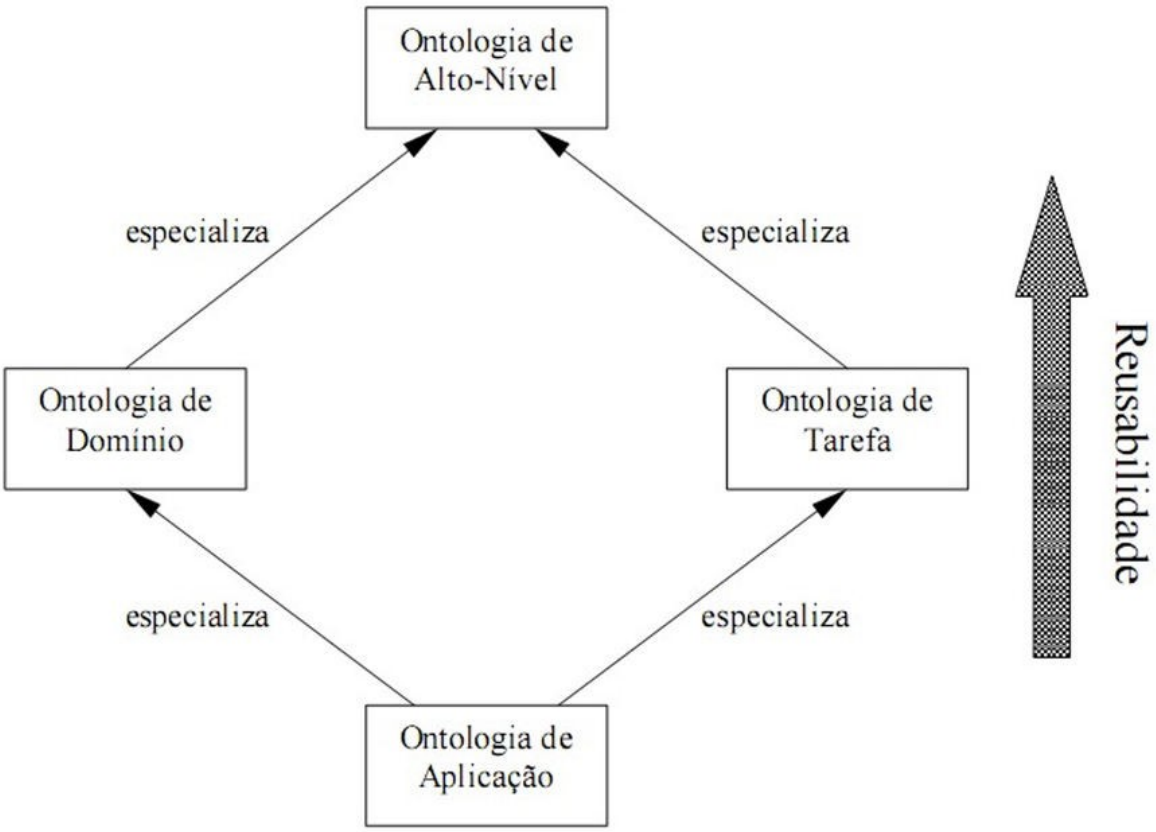


Figura 4.1: Classificação das Ontologias - Fonte Guarino 1998

GUARINO (June 1998) propõe quatro tipos de ontologias, sendo o primeiro deles as ontologias de alto nível, nas quais são descritos conceitos genéricos como espaço, tempo e evento. Estes conceitos são independentes de problemas ou domínios específicos, sendo comum encontrar amplas comunidades de usuários que compartilham ontologias de alto nível.

O segundo tipo são as ontologias de domínio, que descrevem um vocabulário relacionado a um domínio genérico, especializando conceitos presentes nas ontologias de alto nível.

As ontologias de tarefa constituem o terceiro tipo, apresentando uma concepção relacionada a uma tarefa ou atividade genérica.

Por último, as ontologias de aplicação representam o quarto tipo, sendo mais específicas e utilizadas dentro de aplicações específicas. Essa categorização proporciona uma abordagem abrangente para a aplicação eficaz das ontologias em diversos contextos, desde conceitos genéricos até aplicações específicas e direcionadas.

Observa-se que as ontologias de alto nível possuem uma maior capacidade de reuso, uma vez que definem conceitos genéricos, enquanto as ontologias de aplicação apresentam uma menor capacidade de reuso, ao definirem conceitos relacionados a uma aplicação específica. Em todas as fases da construção de uma ontologia, é crucial a utilização de critérios na estruturação do código de linguagem de programação, especialmente na distinção ou comparação das linguagens utilizadas no desenvolvimento.

4.2.1 Aprimoramento Semântico de Dados por Meio de Ontologias

A ontologia pode oferecer contribuições significativas na interpretação adequada de termos como "desinteligências", frequentemente encontrados em bancos de dados policiais, dentro do contexto da análise preditiva criminal. Ontologias, ao fornecerem uma estrutura formal para a representação de conceitos e suas inter-relações, possibilitam uma compreensão mais precisa e consistente dos dados, especialmente em domínios complexos como o da criminologia.

Por exemplo, o termo "desinteligências"pode se referir a informações errôneas ou incompletas registradas em relatórios policiais. Uma ontologia pode definir "desinteligências"como um subconjunto de "dados de inteligência", especificando suas características e relações com outros conceitos GUARINO (June 1998), como "tipos de crime"ou "fontes de informação". Isso não apenas clarifica o significado do termo, mas também permite que algoritmos de análise preditiva tratem essas informações de maneira mais eficaz, evitando confusões que poderiam comprometer a qualidade das previsões.

Ademais, ao integrar ontologias no processamento de dados policiais, é possível harmonizar informações provenientes de diversas fontes, como relatórios de ocorrência, dados de vigilância e informações comunitárias. Essa integração é essencial para a criação de modelos preditivos que sejam não apenas robustos, mas também fundamentados em dados coerentes e bem interpretados Uschold e Gruninger (1996)). Por exemplo, um sistema de análise preditiva pode usar uma ontologia para categorizar dados de crime e suas inter-relações, permitindo a identificação de padrões como o aumento de "desinteligências"associadas a determinados tipos de crime em áreas específicas.

Nesse contexto, a utilização de ontologias, além de facilitar a interpretação

correta de termos e dados, pode também potencializar a eficácia dos sistemas de previsão criminal.

Em todas as etapas de elaboração de uma ontologia, é fundamental aplicar certos critérios na organização do código da linguagem de programação, especialmente ao diferenciar ou comparar as linguagens utilizadas para o desenvolvimento. No contexto da ontologia, as linguagens de programação funcionam como códigos que são interpretados em páginas web, sendo processados ou compilados para que agentes eletrônicos possam acessar as informações GUARINO (June 1998).

Para que a Web 4.0 alcançasse um desempenho superior e deixasse de ser meramente uma web de documentos estáticos, foram propostas diversas tecnologias, entre elas as ontologias. De acordo com Piteri (Piteri e Rodrigues, 2011), a intenção com essa tecnologia foi *".. atribuir sentido e significado ao conteúdo de documentos, atuando como ferramenta de representação do conhecimento. "*.

Por conta da eficiente interoperabilidade e representação semântica de dados, pode ser melhorado o gerenciamento de informação. Nesse contexto, o Resource Description Framework (RDF) surge como uma linguagem de modelagem fundamental, proporcionando uma estrutura padrão para representar informações e suas relações na forma de triplas sujeito-predicado-objeto. A abordagem baseada em grafos do RDF permite uma representação semântica robusta e flexível.

As triplas RDF, compostas por *"sujeito-predicado-objeto"* (Rodrigues e Maciel, 2022), descrevem relações entre entidades e valores. A definição dos aspectos criminais a serem representados é crucial, uma vez que a precisão e a abrangência dessas informações impactam a eficácia das análises e interpretações subsequentes. Essa escolha deve considerar não apenas o crime em si, mas também o contexto social, legal e geográfico em que ocorreu (Ishak, 2022). Além disso, é imperativo ponderar sobre a sensibilidade de certas informações, garantindo a preservação da privacidade das vítimas e demais envolvidos, ao mesmo tempo em que se proporciona uma visão holística e compreensiva do incidente para fins de investigação e políticas públicas.

Complementando o RDF, o SPARQL (SPARQL Protocol and RDF Query Language) se destaca como uma linguagem de consulta específica para dados RDF. Essa linguagem oferece meios eficazes para recuperar informações armazenadas em formato RDF, permitindo consultas complexas e detalhadas. A utilização do SPARQL facilita a extração pr ecisa de dados semânticos, contribuindo para a eficiência na recuperação e análise de informações.

Por fim, o XML (Extensible Markup Language) desempenha um papel crucial na representação estruturada e interoperável de dados. Com sua capaci-

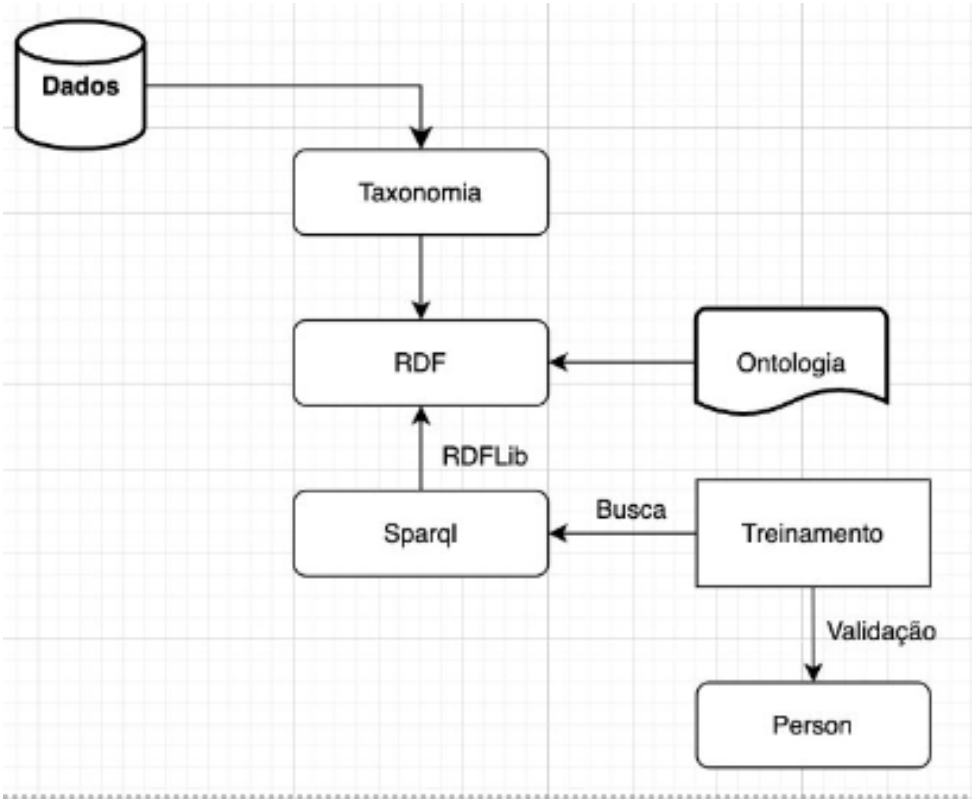


Figura 4.2: RDF - Fonte adaptada de Biz, Bettoni, Thomaz, Santos e Pavan (2014)

dade de definir esquemas personalizados, o XML facilita a troca de informações entre sistemas heterogêneos. Sua flexibilidade e adaptabilidade o tornam uma escolha frequente em contextos nos quais a padronização e a hierarquia de dados são fundamentais.

Em conjunto, RDF, SPARQL e XML formam uma tríade muito eficiente, gerando uma base sólida para a representação semântica, consulta ágil e troca de dados estruturados em ambientes computacionais complexos.

4.2.2 Aplicação da Ontologia na Predição Criminal

No processo de predição criminal, a ontologia pode ser usada para padronizar e esclarecer conceitos que podem variar significativamente entre diferentes fontes de dados ou agentes envolvidos. Por exemplo, a definição de crimes como "roubo"pode diferir entre jurisdições ou contextos, levando a inconsistências nos dados. Uma ontologia pode unificar essas definições, garantindo que o termo "roubo"seja interpretado de maneira consistente em todos os contextos, o que é essencial para a precisão das previsões.

Tabela 4.1: Definição de Crimes por Jurisdição

Conceito	**Definição Jurisdição A**	**Definição Jurisdição B**
Agressão	Contato físico não consensual	Contato físico com dano físico visível

4.2.3 Melhoria na Previsão através da Semântica

A ontologia também é essencial na definição de variáveis que influenciam as previsões criminais. Por exemplo, ao considerar o impacto das condições climáticas na atividade criminal, termos como "frio"podem ter significados diferentes. Em uma região, "frio"pode ser definido como temperaturas abaixo de $17\,^{\circ}\mathrm{C}$, enquanto em outra, pode ser abaixo de $11\,^{\circ}\mathrm{C}$. A ontologia pode padronizar esses conceitos, assegurando que as análises preditivas levem em conta as variáveis contextuais de forma precisa e contextualizada.

Tabela 4.2: Definição de Frio por Região

Região	**Definição de Frio**
Região A	Temperatura abaixo de 17°C
Região B	Temperatura abaixo de 11°C

Essa padronização é particularmente importante em modelos preditivos que utilizam múltiplas fontes de dados com diferentes critérios de registro. Por exemplo, ao integrar dados de incidentes criminais com informações meteorológicas, a ontologia garante que o termo "frio"seja interpretado corretamente em todas as variáveis analisadas.

4.2.4 Expansão da Ontologia para Outras Modalidades Criminais

A ontologia pode ser estendida para lidar com uma ampla variedade de crimes e suas características específicas. Por exemplo, ao considerar crimes como "fraude"ou "cybercrime", a ontologia pode definir claramente os diferentes tipos de fraude (ex.: fraude financeira, fraude de identidade) e suas características, facilitando a análise preditiva.

Tabela 4.3: Tipos de Fraude e Características

Tipo de Fraude	**Características Principais**
Fraude Financeira	Envolvimento de transações financeiras falsas ou enganosas
Fraude de Identidade	Uso não autorizado de dados pessoais para ganho financeiro

4.2.5 Integração de Ontologias no Contexto das Variáveis Preditoras

No contexto de variáveis que influenciam a predição criminal, a ontologia ajuda a definir e padronizar conceitos como "período do dia"(ex.: manhã, tarde, noite), "densidade populacional"e "áreas de alta criminalidade". Por exemplo, ao modelar o impacto do horário do dia na ocorrência de crimes, a ontologia

pode definir claramente os intervalos de tempo para "noite"e "madrugada", que podem variar entre regiões.

Tabela 4.4: Definição de Período do Dia por Região

Período do Dia	**Definição Região A**	**Definição Região B**
Noite	18:00 - 22:00	20:00 - 24:00
Madrugada	22:00 - 06:00	00:00 - 06:00

4.2.6 Exemplos de RDF para Representação de Conceitos

Para ilustrar como a ontologia pode ser implementada na prática, considere os seguintes exemplos de RDF (Resource Description Framework), que descrevem relações entre crimes, suas características e variáveis contextuais:

```
@prefix ex: <http://example.org/crime#> .

ex:Assault a ex:Crime ;
    ex:hasDefinition "Physical contact without consent" ;
    ex:hasSeverity ex:High ;
    ex:occursIn ex:LocationA .

ex:Fraud a ex:Crime ;
    ex:hasType ex:FinancialFraud ;
    ex:involves ex:MonetaryTransaction ;
    ex:hasSeverity ex:Moderate ;
    ex:occursIn ex:LocationB .

ex:WeatherCondition a ex:ContextualVariable ;
    ex:hasType ex:Temperature ;
    ex:hasThreshold "Below 17 degrees Celsius" ;
    ex:relatedTo ex:CrimeRateIncrease .

ex:TimeOfDay a ex:ContextualVariable ;
    ex:hasPeriod ex:Night ;
    ex:occursFrom "18:00" ;
    ex:occursTo "22:00" ;
    ex:relatedTo ex:IncreasedAssaultRisk .
```

Esses exemplos RDF demonstram como conceitos relacionados a crimes e suas variáveis contextuais podem ser formalmente definidos e inter-relacionados para facilitar análises preditivas mais precisas.

4.2.7 Conclusão

A utilização de ontologias no contexto da predição criminal representa um grande avanço na forma como os dados são interpretados e analisados. Ao fornecer uma estrutura semântica consistente, as ontologias permitem que conceitos complexos, como definições de crimes e variáveis contextuais, sejam padronizados e compreendidos de maneira uniforme, independentemente da diversidade de fontes e jurisdições. Essa padronização é muito importante para a construção de modelos preditivos robustos, capazes de identificar padrões e tendências com maior precisão.

As ontologias também facilitam a integração de dados heterogêneos, como registros de ocorrências policiais, informações meteorológicas e dados socioeconômicos, proporcionando uma visão holística e interconectada dos fatores que influenciam a criminalidade. Ao formalizar as relações entre conceitos e variáveis, as ontologias não apenas melhoram a qualidade das previsões, mas também contribuem para uma tomada de decisão mais informada e eficiente na área de segurança pública.

4.3 Redes Neurais Artificiais

As redes neurais artificiais são um conjunto de algoritmos de aprendizado supervisionado empregados para enfrentar diversos desafios em Inteligência Artificial. Sua concepção é inspirada nas redes de neurônios do cérebro de animais, que aprendem por meio da experiência.

Redes Neurais Artificiais (RNA) são estruturas computacionais inspiradas nas redes neurais biológicas. As RNAs são compostas por nós, chamados de neurônios, e cada neurônio é conectado a outro através de conexões. Para cada conexão existe um valor, conhecido como peso Kasabov (1996) Olligschlaeger (1997). As Redes Neurais são conhecidas como modelos conexionistas devido a essas características. O aprendizado de uma RNA é obtido através de algoritmos cujo objetivo é modificar os pesos para que a rede possa gerar resultados de acordo com os exemplos de treinamento. A área da ciência que aplica as RNAs para o processamento de informações é conhecida como Neurocomputação.

A motivação para o desenvolvimento desses algoritmos advém do fato de que os computadores são extremamente rápidos e eficientes em áreas que exigem cálculos matemáticos complexos, tarefas que levariam muito tempo para os seres humanos realizarem. No entanto, os computadores demonstram ineficiência no reconhecimento de padrões, especialmente aqueles obtidos por meio da visão Braga et al. (2000).

Estas redes, em virtude de sua origem, compartilham várias características com o sistema nervoso humano:

- A menor unidade de processamento de informações é um conjunto de elementos simples denominados neurônios; - Esses neurônios podem receber e enviar estímulos para outros neurônios ou para o ambiente; - Informações e sinais são transmitidos entre neurônios através de conexões denominadas sinapses; - A eficiência de uma sinapse, representada por um peso ou força associada, corresponde à informação armazenada no neurônio e, por conseguinte, na rede; - O conhecimento é adquirido do ambiente por meio de um processo conhecido como aprendizado, que é essencialmente responsável por adaptar as forças de conexão, ou valores de peso no caso artificial, aos estímulos externos.

Portanto, o aprendizado consiste em encontrar as forças de conexões apropriadas para produzir padrões de ativação satisfatórios em determinadas circunstânciasCalis (2018).

4.3.1 Neurônio Artificial

O funcionamento de um neurônio biológico ocorre da seguinte forma: o estímulo é recebido por canais localizados nas sinapses, permitindo que os íons fluam para dentro e para fora do neurônio. Um potencial de membrana surge como resultado da integração das entradas neurais, determinando se um dado neurônio produzirá ou não um pico (potencial de ação). Esse pico faz com que os neurotransmissores sejam liberados no final do axônio, formando sinapses com os dendritos de outros neurônios. O potencial de ação só ocorre quando o potencial de membrana ultrapassa um nível de limiar crítico. Diferentes entradas podem fornecer quantidades variadas de ativação, dependendo da quantidade de neurotransmissor liberado pelo emissor e de quantos canais no neurônio pós-sináptico são abertos Haykin (2001). Uma ilustração de um neurônio biológico pode ser vista na Figura 4.3 – suas principais partes citadas são apontadas.

A partir desta representação, sucessivas melhorias no modelo de um neurônio artificial levaram a representação generalizada na Figura seguinte:

Para cada um dos k neurônios, as entradas $x_1, x_2, \ldots, x_m$ são multiplicadas pelos respectivos pesos $w_{k1}, w_{k2}, \ldots, w_{km}$ e, em seguida, somadas na junção somatória. Essa soma, denotada por u_k, é então passada como entrada para uma função de ativação definida pelo projetista da rede. Esta função produz um valor específico que representa a saída final y do neurônio.

Um neurônio artificial também pode ser representado por uma função matemática, como descrito na Equação:

A saída y_k de um neurônio é o resultado de sua função de ativação, que recebe como entrada um valor u_k. Este valor u_k é o somatório das entradas multiplicadas por seus respectivos pesos.

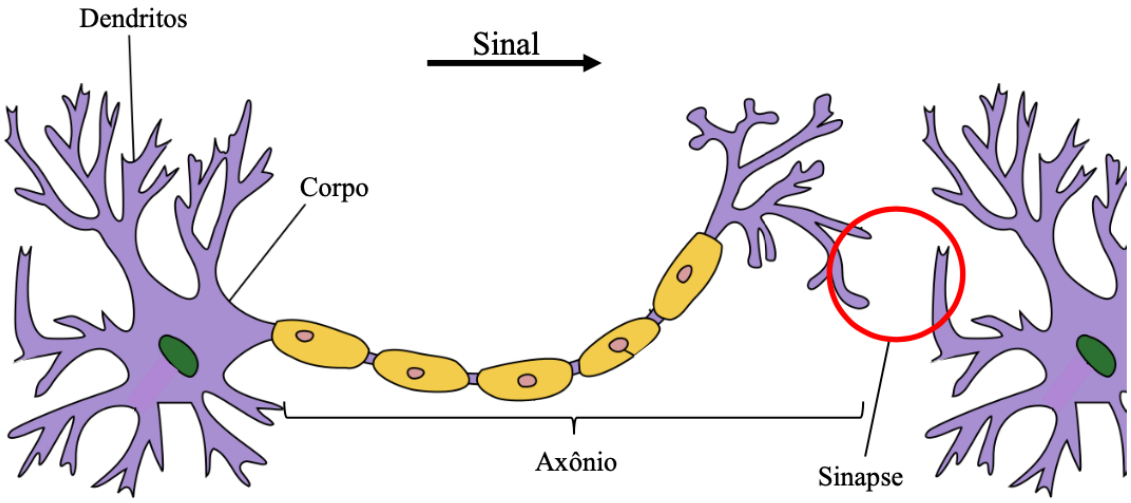


Figura 4.3: Neurônio Biológico - Fonte adaptada de https://commons.wikimedia.org/w/index.php?curid=7616130

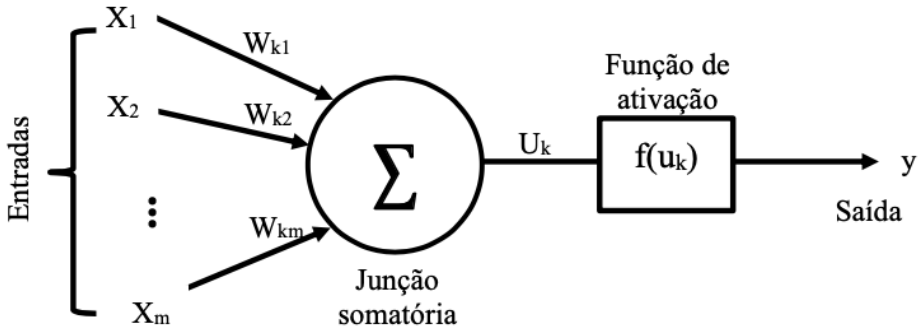


Figura 4.4: Representação genérica de um neurônio artificial, Fonte adap. de Dewinter et al. (2020)

$$y_k = f(u_k) = f\left(\sum_{j=0}^{m} w_{kj} x_j\right)$$

Figura 4.5: Função matemática de um neurônio artificial, Fonte adap. de Dewinter et al. (2020)

4.3.2 Arquitetura e aprendizado de redes neurais

O funcionamento do sistema nervoso dos animais ainda não é completamente compreendido, e os detalhes sobre como as interconexões entre os neurônios são formadas permanecem em grande parte desconhecidos. No entanto, é sabido que uma grande quantidade de neurônios interage para realizar as tarefas cognitivas do indivíduo. Certamente, um único neurônio isolado não seria capaz de realizar tarefas complexas Haykin (2001).

A característica mais notável das redes neurais é sua capacidade de aprender. O aprendizado em uma rede neural ocorre através da alteração dos pesos nas conexões entre os neurônios. Cada exemplo de treinamento x_i é apresentado à rede durante o processo de aprendizado, e os pesos são ajustados para que a saída da rede corresponda ao valor desejado y_i associado ao exemplo x_i. O processo de aprendizado envolve a aplicação de todo o conjunto de exemplos de treinamento X e a subsequente modificação dos pesos, permitindo à rede neural generalizar a função aprendida. Para avaliar a precisão da rede treinada, são utilizados exemplos de teste, que são exemplos de treinamento que não foram incluídos no processo de aprendizagem da rede.

As Redes Neurais Artificiais (RNAs) aprendem o que é proposto a aprender através de algoritmos de aprendizado, que podem ser classificados da seguinte forma:

- **Aprendizado Supervisionado**: A rede neural aproxima uma função $f(X) = y$ a partir de exemplos que incluem atributos $x_1, \ldots, x_n$ e o resultado desejado y_k. O conhecimento adquirido é "codificado"nos pesos das conexões.

- **Aprendizado Não-Supervisionado**: Em vez de receber um conjunto de atributos X e a saída desejada, a RNA recebe apenas os atributos. As Redes Neurais de Mapeamento Auto-Organizável (SOM, do inglês *Self-Organizing Maps*) são um exemplo de rede com aprendizado não supervisionado.

- **Aprendizado por Reforço**: Nesse tipo de aprendizado, a rede é recompensada quando a saída é considerada satisfatória, o que aumenta o valor dos pesos das conexões. Se a saída for considerada insatisfatória, a rede é penalizada e os pesos das conexões envolvidas são reduzidos, diminuindo seu impacto no resultado da rede.

Por essa razão, as redes neurais artificiais são organizadas em conjuntos de neurônios denominados camadas. A primeira camada de uma rede é geralmente chamada de camada de entrada, responsável por receber o estímulo externo. Existem também uma ou mais camadas intermediárias, ou ocultas,

e uma camada de saída, que externa o resultado da rede Braga et al. (2000). Com exceção das redes recorrentes, as redes neurais são, em sua maioria, do tipo feedforward, onde a propagação do sinal ocorre sempre da entrada para a saída Calis (2018).

Como o trabalho usará uma rede com arquitetura *feedforward*, então é oportuno, discutir o algoritmo mais comum associado a essa arquitetura. O algoritmo de aprendizado mais conhecido para Redes Neurais Artificiais (RNAs) com arquitetura feedforward é o *backpropagation* Riedmiller e Braun (1993). A regra Delta é a base para o cálculo da atualização dos pesos utilizada no *backpropagation* Widrow e Lehr (1990). O objetivo do *backpropagation* é minimizar o erro global, expresso por

$$\text{Err} = \sum_{p} \sum_{j} \text{Err}_{jp}, \tag{4.1}$$

onde o erro de um exemplo p pode ser calculado, por exemplo, pelo Erro Médio Quadrático (Mean Square Error, MSE):

$$\text{Err}_{jp} = \frac{\epsilon_j^2}{2}, \tag{4.2}$$

como descrito por Allen (1971). Os pesos das conexões são atualizados com base na regra Delta. Os ajustes definidos por essa regra são dados pela seguinte expressão:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \cdot \epsilon_j(t) \cdot x_i, \tag{4.3}$$

onde:

- i = índice do sinal de entrada,
- j = índice do neurônio da camada de saída,
- t = iteração,
- $w_{ij}(t+1)$ = valor do peso ajustado,
- $w_{ij}(t)$ = valor do peso anterior,
- η = taxa de aprendizado,
- $\epsilon_j(t)$ = valor do erro para o neurônio j, conforme a expressão a seguir,
- x_i = valor de entrada.

O erro para o neurônio j é calculado por:

$$\epsilon_j(t) = d_j(t) - y_j(t), \tag{4.4}$$

onde:

- $d_j(t)$ = saída desejada para o neurônio j,
- $y_j(t)$ = saída calculada para o neurônio j.

Na próxima figura, observa-se a ilustração de uma rede feedforward de múltiplas camadas.

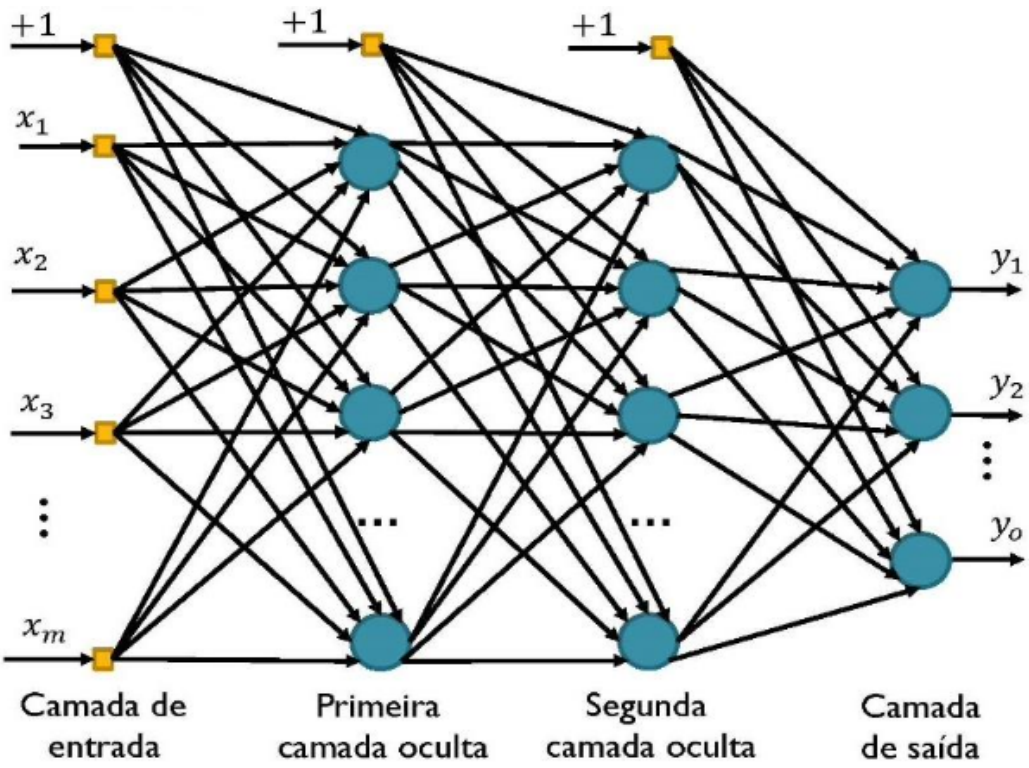


Figura 4.6: Rede neural artificial de múltiplas camadas, Fonte adap. de Dewinter et al. (2020)

Considerando uma rede neural como a apresentada na Figura, o processo de aprendizagem consiste em apresentar os padrões de entrada para a rede, ou seja, os dados de treinamento. Esses dados permitem que a rede altere seus parâmetros livres de modo a produzir padrões de resposta desejados para os dados de entrada. É importante destacar que esses dados iniciais são rotulados, ou seja, sabe-se de antemão a qual classe o exemplo atual pertence, permitindo assim obter medidas de desempenho da rede. Esse processo de aprendizagem, com dados rotulados, é chamado de aprendizagem supervisionada Braga et al. (2000).

O desempenho da rede é medido ao se verificar a taxa de aprendizado, ou seja, a capacidade de generalização da rede Haykin (2001). Os parâmetros livres, em geral, são pesos ajustados por algum algoritmo de aprendizado. Se o treinamento for bem-sucedido, ou seja, a rede é capaz de generalizar o

padrão em questão, novos exemplos daquele padrão podem ser apresentados para que a rede possa agora classificá-los Calis (2018).

4.3.3 Redes Neurais Convolucionais

No campo da aprendizagem de máquina, as Redes Neurais Convolucionais (CNNs) têm se destacado como uma poderosa classe de modelos, especialmente na análise de dados visuais, como imagens. Originadas como uma evolução das redes neurais tradicionais, as CNNs foram desenvolvidas para superar desafios específicos associados ao processamento de informações complexas, como padrões visuais em conjuntos de dados extensos. No campo da aprendizagem de máquina, as Redes Neurais Convolucionais (CNNs) têm se destacado como uma poderosa classe de modelos, especialmente na análise de dados visuais, como imagens. Originadas como uma evolução das redes neurais tradicionais, as CNNs foram desenvolvidas para superar desafios específicos associados ao processamento de informações complexas, como padrões visuais em conjuntos de dados extensos.

Esta subseção abordará o papel das CNNs, destacando sua arquitetura distintiva e eficácia em tarefas que exigem análise detalhada de padrões espaciais. Quando compreendemos a estrutura única das CNNs e sua capacidade de extrair características relevantes, podemos explorar como essas redes desempenham um papel importante na interpretação de dados visuais para aplicações específicas, gerando avanços significativos em diversos domínios, incluindo a segurança pública.

Uma das propriedades mais importantes de uma rede neural artificial é a capacidade de aprender por intermédio de exemplos e fazer inferências sobre o que aprendeu, melhorando gradativamente o seu desempenho. As redes neurais utilizam um algoritmo de aprendizagem cuja tarefa é ajustar os pesos de suas conexões (Braga et al., 2000, Carvalho e Tonet, 1994).

A arquitetura de uma rede neural desempenha seu papel ao determinar a natureza dos problemas nos quais a rede pode ser aplicada. Esta arquitetura é delineada pelo número de camadas, podendo ser uma única camada ou múltiplas camadas, pela quantidade de nós em cada camada, pela natureza da conexão entre esses nós (*feedforward* ou *feedback*) e pela sua topologia geral (Haykin, 2001).

Enquanto nas RNAs tradicionais cada neurônio na camada de entrada é conectado a cada neurônio da camada de saída na camada seguinte, nas CNNs estas camadas, também conhecidas por camadas totalmente conectadas, são as últimas.

O processo de propagação de informações em redes neurais tradicionais ocorre apenas em uma direção, da camada de entrada para a camada de saída, sem retroalimentação. Cada conexão entre neurônios é associada a

um peso, e a rede aprende ajustando esses pesos durante o treinamento. O treinamento é geralmente realizado usando algoritmos de otimização, como o gradiente descendente, com o objetivo de minimizar a diferença entre as saídas da rede e os valores desejados.

Segundo Calis (2018), as camadas cruciais em uma Rede Neural Convolutiva (CNN) são as camadas convolucionais. Essas camadas funcionam como um extrator automático de características das imagens, de forma que as características inferidas pela rede neural são as mais relevantes para o processo de classificação.

Os parâmetros de entrada dessa camada são a quantidade de filtros, que podem ser aprendidos, bem como suas dimensões, o tamanho do passo a ser dado na convolução, isto é, quantos *pixels* serão desconsiderados a cada iteração do operador. Depois de aplicar todos os n filtros ao volume de entrada, tem-se n mapas de ativação bidimensionais, como pode ser visto na Figura - Mapas de ativação empilhados para próxima camada(a). Em seguida, empilham-se os mapas de ativação ao longo da dimensão de profundidade da matriz para formar a saída final na Figura - Mapas de ativação empilhados para próxima camada(b).

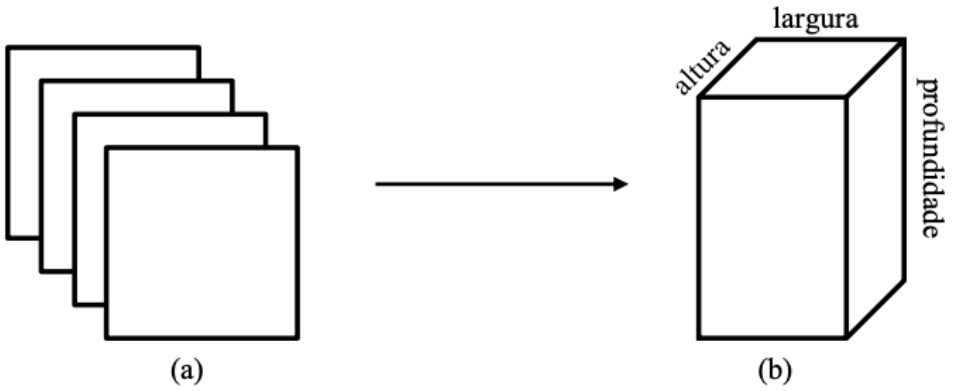


Figura 4.7: Mapas de ativação empilhados para próxima camada, Fonte: (Couto, 2023)

Cada entrada no volume de saída de um modelo de aprendizagem de máquina para predição criminal representa a saída de um neurônio que analisa uma pequena região dos dados geoespaciais. Dessa forma, o sistema aprende a ativar filtros quando identifica padrões específicos em determinadas localizações no mapa.

Nas camadas inferiores do modelo, os filtros podem ser ativados ao reconhecer características de baixo nível, como agrupamentos de ocorrências ou áreas de alta densidade criminal. Em contraste, nas camadas mais profundas, os filtros podem ser ativados na presença de padrões de alto nível, como tendências sazonais ou geográficas que indicam risco aumentado, como a proximidade a eventos ou a presença de infraestruturas específicas.

Na abordagem tradicional de análise de dados para predição criminal, a identificação desses padrões exigiria o desenvolvimento de algoritmos específicos para detectar e mapear características relevantes nos dados, o que pode ser otimizado através de redes neurais artificiais Haykin (2001).

Após as convoluções, uma função de ativação não linear, como a ReLU (*Rectified Linear Unit*), é aplicada à saída dessas operações. O processo de convolução é acompanhado por uma combinação de outros tipos de camadas, visando reduzir a largura e a altura do volume de entrada, bem como mitigar o *over fitting*, que representa a perda de capacidade de generalização da rede. Essa sequência de operações continua até atingir o final da rede, onde uma ou duas camadas completamente conectadas (também conhecidas como *fully connected* - FC) são aplicadas para obter as classificações finais de saída.

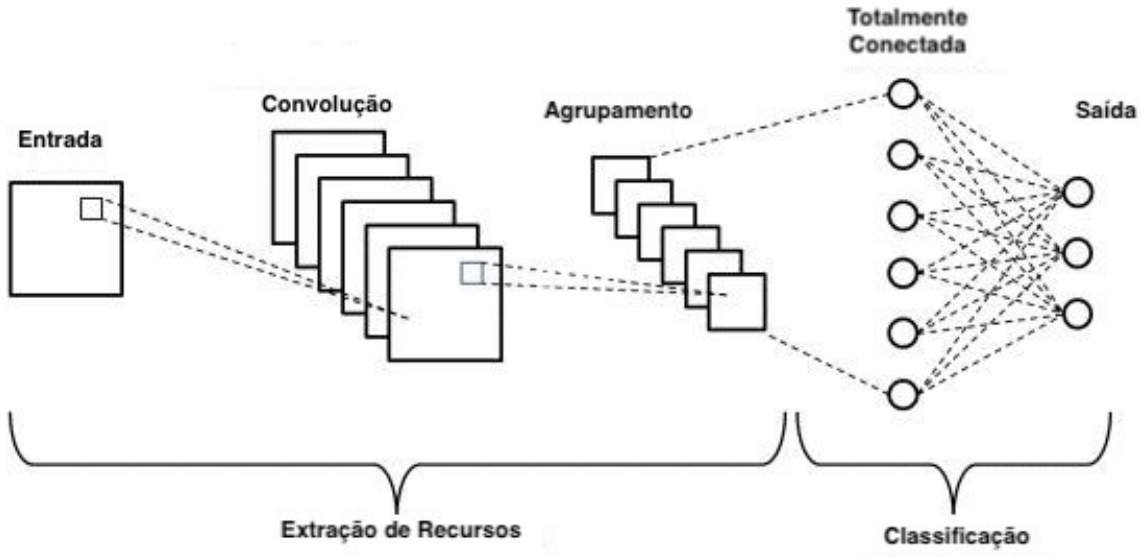


Figura 4.8: Representação de camadas de uma CNN, Fonte: (Couto, 2023)

Cada camada em uma CNN aplica um conjunto distinto de filtros, geralmente em número de centenas ou milhares, e combina os resultados, alimentando a saída na próxima camada da rede. Durante o treinamento, a CNN aprimora automaticamente os valores desses filtros, seguindo uma abordagem semelhante às redes neurais tradicionais.

A diferenciação fundamental entre as CNNs e as redes neurais tradicionais reside na disposição das camadas, com as camadas convolucionais desempenhando um papel central na extração de características relevantes.

Quando percebemos a importância das camadas convolucionais, que aplicam filtros a regiões locais da imagem, as CNNs possibilitam verificar invariância local e composicionalidade, dois benefícios muito importantes. A invariância local permite que a rede reconheça objetos independentemente de sua

posição na imagem, enquanto a composicionalidade possibilita a construção de representações de alto nível a partir de características de baixo nível.

Esses benefícios impulsionam a capacidade de classificação de imagens e também abrem caminho para análises preditivas mais sofisticadas no contexto do análise preditiva de crimes. A capacidade de aprender automaticamente padrões complexos e construir representações hierárquicas sugere que as CNNs têm o potencial não apenas de identificar eventos passados, mas também de antecipar e responder a cenários futuros. Sendo assim, ao gerar a análise preditiva podemos perceber que as CNNs podem ampliar significativamente a eficácia das operações policiais, gerando uma abordagem proativa para a segurança pública.

4.4 Análise de Dados

Os tipos de análises de dados são fundamentais para explorar e utilizar informações de forma estratégica em diversas áreas. A análise descritiva foca em descrever padrões e características atuais dos dados, utilizando estatísticas descritivas e visualizações gráficas. Já a análise diagnóstica busca entender as causas raiz de eventos passados, examinando relações causais nos dados por meio de técnicas como análise de regressão e correlação. Por outro lado, a análise preditiva utiliza modelos estatísticos e algoritmos de *machine learning* para prever eventos futuros com base em dados históricos, sendo aplicável em previsões de mercado e tendências. Complementando, a análise prescritiva vai além da previsão, recomendando ações específicas para otimizar resultados futuros, como sistemas de recomendação e otimização de recursos. Esses tipos de análises orientam decisões estratégicas em diversas disciplinas, permitindo desde a compreensão do presente até a antecipação e influência sobre o futuro com base em dados disponíveis. No contexto da análise de dados, há uma progressão significativa de complexidade e valor agregado à medida que avançamos dos tipos descritivo e diagnóstico para os preditivo e prescritivo. As análises descritiva e diagnóstica são essenciais para entender o que aconteceu e por quê, oferecendo uma visão retrospectiva dos dados, como uma fotografia do passado. Em contraste, a análise preditiva busca prever eventos futuros com base nos dados disponíveis, permitindo não apenas antecipar eventos, mas também sugerir ações estratégicas. É neste ponto que entra a análise prescritiva, que não apenas prevê o que pode acontecer, mas também recomenda decisões para influenciar proativamente o futuro, nesse trabalho focaremos na análise preditiva, preparando terreno para trabalhos futuros na área de análise prescritiva.

Essas formas de análise que lidam com o futuro são categorizadas como Analytics Avançado, sendo utilizadas com precisão por poucas empresas, como

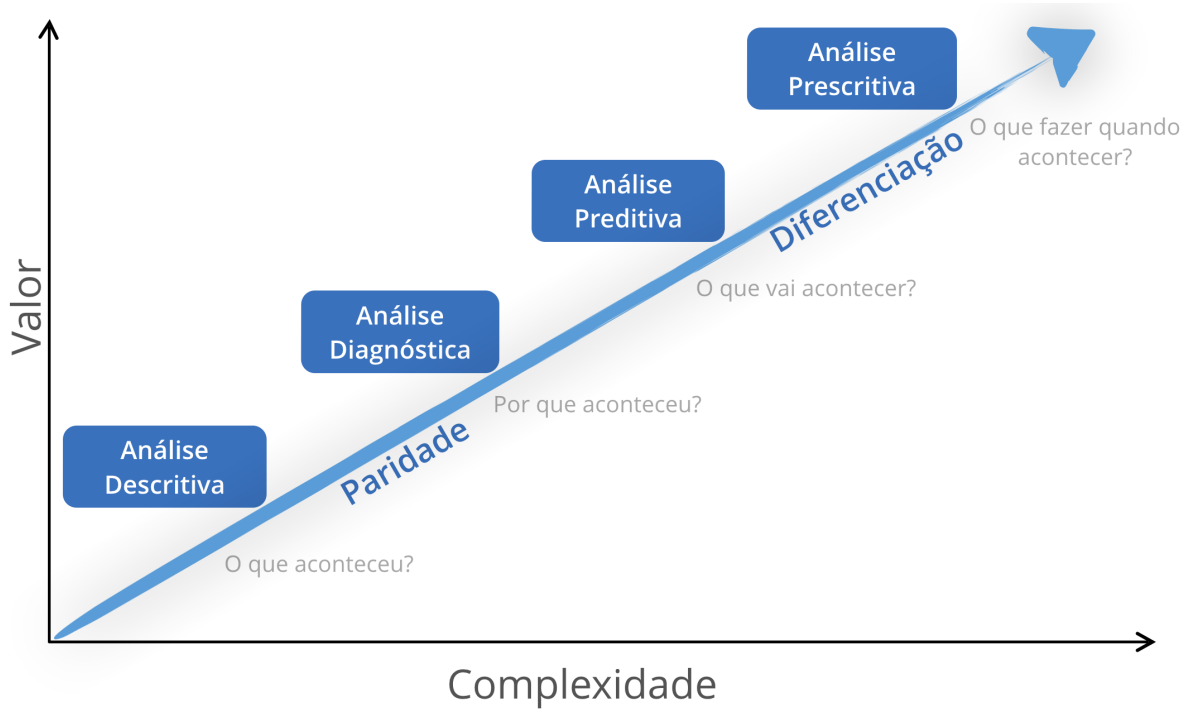


Figura 4.9: Análise de Valor x complexidade - Fonte: Wikipedia (Jun/2024)

exemplos notáveis a Netflix e o Spotify. Enquanto as análises descritiva e diagnóstica são comumente associadas a técnicas de Data Warehousing e Business Intelligence (DW/BI), ferramentas como o Google Analytics também desempenham esses papéis, mesmo não sendo exclusivamente um DW/BI, destacando a diversidade de aplicações e capacidades analíticas no contexto moderno de dados.

Dentro das iniciativas voltadas para assegurar a segurança pública, a atividade de policiamento, por meio da presença ostensiva, destaca-se como a faceta mais visível da resposta da sociedade ao crime e à preservação da ordem pública. Essa prática é essencial como uma medida de controle social, necessária para promover o bem-estar e a convivência harmoniosa (Cordner et al., 2007). Uma das características fundamentais desse policiamento é a autorização coletiva da sociedade para o uso efetivo da força pelos agentes públicos, quando alguém viola as normas de convívio social. (Bayley, 2002)

Assim, os modelos tradicionais de policiamento buscam prevenir e reagir às ocorrências, evitando que o crime ocorra ou restabelecendo os direitos violados, como a integridade física e o patrimônio, diante de uma agressão real ou iminente. Nesse sentido, é essencial que os criminosos não cometam ilícitos na presença policial, sendo crucial a alocação eficiente dos recursos disponíveis para evitar delitos. Isso envolve a presença policial em locais propícios à ocorrência de crimes (Koper, 1995, Hart e Zandbergen, 2014, Dewinter et al., 2020), com a alocação estratégica de recursos, como postos policiais ou

viaturas, sendo objeto de estudo por décadas.

Com o objetivo de otimizar a alocação dos recursos policiais, seja na disposição de postos policiais ou viaturas, seja na definição de roteirização, busca-se empregar esses recursos nos locais onde podem prevenir crimes ou responder de forma mais rápida a situações de emergência. A busca por soluções de otimização na alocação dos recursos policiais, utilizando métodos como o Problema de Localização de Facilidades e o Problema de Roteamento de Veículos, tem sido uma área de estudo contínua ao longo das décadas. (Simpson e Hancock, 2009, Dewinter et al., 2020)

Nesse contexto, têm sido desenvolvidos métodos de Policiamento Preditivo (*Predictive Policing*) e Policiamento Inteligente (*Smart Policing*) para estabelecer estratégias adequadas de policiamento com foco na redução dos índices criminais. Essas abordagens priorizam ações proativas por meio da eficaz utilização de dados, buscando aprimorar a análise, medir o desempenho e incentivar a inovação (Coldren Jr et al., 2013). O objetivo central desses métodos e estratégias é antecipar a atividade policial em relação aos eventos criminosos, adotando ações preventivas baseadas na análise de dados (Perry, 2013).

Diversos métodos de policiamento têm sido desenvolvidos para identificar principalmente onde e quando ocorrerão eventos criminais. A alocação eficiente de recursos nos locais de maior risco visa evitar esses "crimes previstos", considerando as limitações dos recursos disponíveis (Perry, 2013, Hart e Zandbergen, 2014, Moses e Chan, 2018). A capacidade de "prever"ou "predizer"crimes parte do reconhecimento de que os crimes não ocorrem de forma homogênea no espaço geográfico, concentrando-se em determinadas áreas e locais conhecidos como *hotspots*. Essa concentração pode ser explicada pela interação entre vítima e infrator, bem como pelas oportunidades existentes para a prática de crimes (Chainey et al., 2008, Ratcliffe, 2010, Hart e Zandbergen, 2014, Rummens et al., 2017, Braga et al., 2000).

Sendo assim o Policiamento Preditivo pode ser definido como a "aplicação de técnicas analíticas - principalmente técnicas quantitativas - para identificar alvos prováveis para intervenção policial e prevenir crimes ou resolver crimes passados através de previsões estatísticas"(Perry, 2013). Em síntese, o policiamento preditivo representa um método no processo de tomada de decisão dos gestores de segurança pública, utilizando técnicas estatísticas e preditivas no planejamento de ações policiais proativas nos locais de maior risco, buscando alcançar maior eficiência (melhor emprego dos recursos), eficácia (atingindo os resultados esperados) e efetividade (gerando impacto no local de atuação).

Nesse contexto, uma modelagem de policiamento inteligente deve considerar a análise para identificação das características criminais no ambiente e no tempo, antecipando seu acontecimento. No entanto, esse processo de-

pende fundamentalmente do planejamento de ações policiais baseado nessas informações, levando em conta a otimização dos recursos.

4.4.1 Principais Tecnologias Relacionadas à Análise Preditiva

Embora o principal objetivo dos modelos de predição seja, em geral, alcançar a melhor performance possível, algumas aplicações exigem um certo nível de interpretabilidade. A interpretabilidade permite um melhor entendimento do funcionamento do modelo, facilitando melhorias e correções. Além disso, compreender o funcionamento do modelo ajuda a identificar as características mais importantes, proporcionando um entendimento mais profundo do problema de Oliveira Caires (2022). Isso é relevante tanto para modelos de classificação quanto para os de regressão.

Os modelos gerados com Floresta de Decisão são reconhecidos por sua alta interpretabilidade, pois permitem que um ser humano compreenda facilmente quais atributos influenciaram a tomada de decisão ao observar a estrutura da árvore. Em contraste, os algoritmos Floresta Aleatória e Gradient Boosting, embora também baseados em Árvores de Decisão, não oferecem a mesma facilidade de visualização. Isso ocorre porque esses algoritmos geralmente consistem em centenas de árvores, tornando-se modelos do tipo "caixa preta".

Para proporcionar algum nível de interpretabilidade a esses modelos, podem ser empregadas técnicas que estimam a importância dos atributos no processo decisório. A importância de atributos, ou *feature importance*, é uma técnica amplamente utilizada em modelos de ensemble. Ela já vem implementada na maioria das bibliotecas de aprendizado de máquina, como Scikit-learn e LightGBM de Oliveira Caires (2022).

Nos modelos baseados em Árvores de Decisão, a importância dos atributos pode ser calculada pela frequência com que um atributo é utilizado nas diferentes árvores do modelo. Essa métrica reflete a importância do atributo, partindo do pressuposto de que quanto mais vezes a variável é utilizada, maior é o seu impacto na predição de Oliveira Caires (2022).

Python é uma linguagem de programação interpretada e orientada a objetos, que suporta múltiplos paradigmas de programação, incluindo os paradigmas funcional e procedural Python Software Foundation (2021). A linguagem oferece uma ampla gama de funcionalidades, como módulos, exceções, tipagem dinâmica, tipos de dados de alto nível e classes Python Software Foundation (2021). Python é uma linguagem bem estabelecida e uma das mais populares para a computação científica Pedregosa et al. (2011).

A popularidade de Python no campo da computação científica e Aprendizado de Máquina se deve, em grande parte, à vasta coleção de bibliotecas e *frameworks* disponíveis. Algumas das bibliotecas mais utilizadas incluem:

- **NumPy**: Fundamental para operações numéricas e manipulação de *arrays*.
- **Pandas**: Essencial para a análise e manipulação de dados.
- **Scikit-learn**: Oferece uma variedade de algoritmos de Aprendizado de Máquina para tarefas de classificação, regressão e *clustering*.
- **TensorFlow e Keras**: Utilizadas para construção e treinamento de redes neurais profundas.
- **Matplotlib e Seaborn**: Ferramentas poderosas para visualização de dados.

Nesse contexto se destacando a Scikit-learn que é um *framework* para Python que integra uma ampla gama de algoritmos de Aprendizado de Máquina de última geração, destinados a problemas de média escala em aprendizado supervisionado e não supervisionado Pedregosa et al. (2011). Scikit-learn oferece métodos para gerar, treinar e avaliar modelos de aprendizado de máquina, utilizando uma linguagem de alto nível com foco na usabilidade e performance.

4.5 Considerações finais

Nesse capítulo foram destacadas tecnologias de Ontologia, Redes Neurais Convolucionais e Análise Preditiva. O capítulo explora como essas tecnologias se entrelaçam para criar um paradigma inovador capaz de apresentar de maneira inédita a compreensão e o aproveitamento de dados complexos.

As Redes Neurais Convolucionais são apresentadas como ferramentas de processamento de informações visuais com potencial extraordinário, capazes de identificar padrões complexos em imagens e dados visuais. A capacidade dessas redes para identificar padrões complexos revoluciona a análise de informações e também inaugura novos horizontes na compreensão de contextos visuais, como os encontrados em vigilância urbana e reconhecimento de padrões comportamentais.

A Ontologia é apresentada como uma ferramenta que proporciona uma compreensão semântica aprimorada dos dados, facilitando a interoperabilidade entre diferentes fontes de informação e melhorando a tomada de decisões.

Por fim, a Análise Preditiva é apresentada como uma ferramenta que possibilita não apenas a resposta reativa a eventos, mas também a antecipação de cenários futuros, permitindo um planejamento estratégico mais eficiente.

Em conjunto, essas tecnologias oferecem uma perspectiva muito promissora para aprimorar a resposta policial, prevenir atividades criminosas e otimi-

zar a alocação de recursos. Contudo, reconhece-se a necessidade de abordar desafios éticos, garantir a privacidade dos dados e avaliar a aceitação social dessas tecnologias no contexto do policiamento.

CAPÍTULO 5

Predição Criminal Utilizando Técnicas de Machine Learning e Análise Espacial

5.1 Considerações Iniciais

A predição criminal é uma área de estudo que visa prever a ocorrência de crimes com base em dados históricos. O objetivo é identificar padrões e tendências que possam ajudar na alocação de recursos de segurança pública de maneira mais eficiente. Neste capítulo, utilizamos técnicas de *machine learning* e análise espacial para prever *hotspots* de crimes em São Paulo.

Neste capítulo é apresentado o contexto e a evolução da previsão empírica para a aplicação de métodos e algoritmos de Aprendizado de Máquina na pesquisa de policiamento inteligente, buscando trabalhar de forma eficaz para lidar com os desafios complexos da segurança pública.

5.2 Previsão Empírica

Todas as corporações, tanto empresariais quanto governamentais, baseiam seu planejamento em dados qualitativos ou quantitativos coletados no dia a dia. Esses dados devem refletir as experiências vividas e demonstrar os erros e acertos cometidos. O planejamento está cada vez mais associado ao uso de sistemas computacionais e técnicas de análise de dados. Modelos de previsão quantitativa utilizam dados históricos para identificar padrões e projetá-los para o futuro. Esses modelos empregam técnicas computacionais e estatísticas para representar e executar ações com base nas informações disponíveis. Portanto, a aquisição de ferramentas de previsão deve ser considerada um di-

ferencial organizacional, pois proporciona suporte adicional às decisões dos gestores. Diversas áreas utilizam previsões para apoiar decisões, como a previsão de preços de ações, pontuação de jogos de futebol, previsão de eventos médicos, ataques a redes de computadores e assaltos residenciais.

A previsão pode ser definida como a obtenção de uma resposta precisa sobre um evento futuro com base em dados passados. O futuro aqui se refere a cenários ou situações que ainda não foram vividos pela organização ou metas a serem alcançadas. Portanto, as previsões devem ser realizadas com base em variáveis independentes, utilizando dados do presente e do passado armazenados nas bases da organização, bem como na experiência dos gestores e outros profissionais envolvidos. As previsões para este trabalho terão um nível considerável de detalhamento para um horizonte temporal curto.

De acordo com Makridakis, Wheelwright e McGee (1998), o recente aumento no uso de técnicas de previsão nas organizações pode ser atribuído a:

- O aumento da complexidade das organizações (número de clientes e produtos) e do mercado (mudanças no mercado e na estrutura de demanda), o que complicou as decisões dos gestores ao considerar todos os fatores relacionados ao futuro desenvolvimento da organização.
- A adoção de procedimentos decisórios mais sistemáticos nas organizações, que requerem justificativas explícitas para cada ação tomada. Ter uma previsão formal ajuda a apoiar tais procedimentos.
- O avanço contínuo das técnicas de previsão e suas aplicações, permitindo que não apenas analistas especializados, mas também gerentes e outros tomadores de decisão compreendam e utilizem essas técnicas.

Um desafio importante é o espaço de aplicação para o desenvolvimento de teorias e sua aplicação correta. Técnicas de previsão só podem ser aplicadas efetivamente em organizações que adotam políticas de planejamento e decisões baseadas em planejamento. Assim, a aplicação das técnicas frequentemente requer adaptações teóricas, o que implica trabalho adicional para ajustar as metodologias. Portanto, é necessário resolver vários problemas antes da aplicação dos métodos preditivos.

Um processo de mineração deve ser seguido para facilitar o desenvolvimento e a aplicação de adaptações e minimizar os problemas decorrentes da execução das técnicas de previsão (Armstrong, 1988, Deroeck, 1991, Mahmoud e Brown, 1992). Esse processo deve conectar a teoria presente na literatura aos problemas encontrados na aplicação prática. O objetivo deste trabalho é desenvolver um processo para a aplicação de previsões de níveis

criminais em qualquer município brasileiro, utilizando algoritmos de aprendizado de máquina. O estudo de caso será a previsão de níveis criminais em áreas demográficas da Região Metropolitana de Fortaleza, utilizando dados criminais e socioeconômicos.

Winklhofer desenvolveu um framework para lidar com as questões da aplicação de técnicas de previsão. De acordo com suas pesquisas, poucos autores elaboraram guias detalhados para o desenvolvimento de sistemas de previsão, com a maioria dos estudos focando em técnicas de previsão estatística (Winklhofer et al., 1996). O framework desenvolvido por eles é dividido em três conjuntos de questões: design, seleção/especificação e avaliação. O design aborda o propósito e o tipo de previsão desejada, os recursos necessários, as características dos recursos humanos e o conjunto de dados a ser utilizado. A seleção e especificação envolvem as técnicas de previsão e questões sobre familiaridade com sua seleção e uso. Finalmente, a avaliação foca nos resultados produzidos pela atividade preditiva, incluindo a apresentação dos resultados, a precisão do modelo e as razões para qualquer perda de precisão.

5.3 Aplicação de Métodos e Algoritmos de Aprendizado de Máquina na Pesquisa de Policiamento Inteligente

O uso de métodos e algoritmos de Aprendizado de Máquina na pesquisa de policiamento inteligente está sendo fundamentado em vários motivos substanciais. Esses métodos oferecem a capacidade de analisar grandes volumes de dados de forma eficiente e identificar padrões complexos que podem não ser evidentes por meio de técnicas tradicionais. A aplicação de Aprendizado de Máquina permite uma análise preditiva avançada, melhorando significativamente a eficácia das estratégias de policiamento e a alocação de recursos. Com a integração de algoritmos avançados, é possível prever atividades criminosas com base em dados históricos e em tempo real, possibilitando uma resposta mais ágil e precisa.

A adoção desses métodos é bem vantajosa na identificação de padrões criminais, antecipação de atividades ilícitas e otimização das operações policiais. Além disso, a aplicação de Aprendizado de Máquina contribui para a personalização das abordagens de segurança, garantindo que as estratégias sejam ajustadas às necessidades específicas e condições locais.

A seguir, apresentam-se os principais fundamentos que justificam a escolha dos métodos utilizados neste trabalho:

Os algoritmos de Aprendizado de Máquina destacam-se pela capacidade de processar grandes volumes de dados de maneira eficiente, identificando pa-

drões complexos muitas vezes imperceptíveis aos métodos tradicionais (Bishop, 2006). Essa característica é particularmente valiosa no contexto da segurança pública, pois permite análises mais precisas e aprofundadas, facilitando a identificação de áreas de maior risco e contribuindo para uma alocação mais estratégica dos recursos policiais.

Além disso, com base na análise de dados históricos, esses modelos possibilitam a previsão de incidentes futuros, fornecendo subsídios relevantes para o planejamento estratégico e para a tomada de decisões operacionais (Breiman, 2001). Dessa forma, as forças policiais podem atuar de forma proativa, antecipando-se aos eventos e adotando medidas preventivas que fortalecem a segurança da população.

Outro benefício significativo é a capacidade dos algoritmos de otimizar o planejamento operacional e a roteirização das patrulhas. Por meio da detecção de padrões de atividade criminosa, os modelos podem prever a demanda por serviços policiais em diferentes regiões e horários, além de sugerir estratégias de patrulhamento mais eficazes (Jordan e Mitchell, 2015). Essa abordagem resulta em respostas mais rápidas e eficientes a ocorrências, colaborando diretamente para a redução da criminalidade e o aprimoramento da segurança pública.

5.4 Medições e Análises de Desempenho

Para avaliar a eficácia do sistema proposto de policiamento inteligente, serão adotadas diversas métricas amplamente utilizadas na avaliação de modelos de aprendizado de máquina. Essas medições permitirão uma análise abrangente do desempenho do sistema, tanto em termos preditivos quanto em sua aplicabilidade prática no contexto da segurança pública.

Uma das métricas empregadas será a correlação de Pearson, que permite identificar a existência e o grau de relação linear entre variáveis relevantes para a segurança pública (Pearson, 1895). Por exemplo, será possível verificar a correlação entre o número de patrulhas alocadas em determinada área e a incidência de crimes nesse local, oferecendo indícios importantes para o ajuste tático-operacional das ações policiais.

Outra métrica fundamental será o Erro Absoluto Médio (MAE), utilizado para quantificar a precisão das previsões realizadas pelo modelo em comparação com os valores observados (Willmott e Matsuura, 2005). Essa análise permitirá avaliar o desvio médio entre as estimativas geradas pelo sistema e os dados reais de criminalidade, refletindo o grau de aderência do modelo à realidade prática.

A acurácia será igualmente considerada, permitindo avaliar a proporção de previsões corretas feitas pelo modelo em relação ao total de previsões (U.M,

1996). Um alto índice de acurácia indicará que o sistema é capaz de identificar corretamente os eventos e padrões relacionados à ocorrência de crimes, fortalecendo sua confiabilidade como ferramenta de apoio à decisão.

Adicionalmente, será aplicada a Curva ROC (Receiver Operating Characteristic), que avalia a capacidade do modelo em distinguir entre duas classes distintas, como a ocorrência ou não de um crime (Haykin, 2001). Essa métrica é particularmente útil na análise de classificadores binários, contribuindo para mensurar a sensibilidade e a especificidade do sistema em cenários reais.

Essas análises de desempenho possibilitarão validar a efetividade da proposta, demonstrando seu potencial para aprimorar as estratégias de policiamento e, consequentemente, promover comunidades mais seguras e resilientes. Ainda assim, é importante ressaltar que o uso de tecnologias preditivas em segurança pública exige atenção especial a questões éticas, à proteção da privacidade dos dados e à aceitação social dessas inovações (Basilio et al., 2019). Esses aspectos serão considerados como parte das limitações e recomendações deste estudo.

5.5 Tratamento dos Dados

Os procedimentos completos de pré-processamento e estruturação dos dados utilizados neste estudo, incluindo os trechos de código em Python e operações estatísticas aplicadas, foram deslocados para o Apêndice A, onde é apresentado uma descrição geral do tratamento realizado e a lógica de construção das variáveis preditivas de forma detalhada.Esta escolha visa não sobrecarregar o corpo do texto, mantendo a clareza e fluidez da apresentação dos resultados principais.

CAPÍTULO 6

Conclusão

O sucesso das corporações, sejam públicas ou privadas, depende basicamente de um planejamento eficaz, que deve considerar o horizonte temporal das decisões: curto, médio e longo prazo.

Tabela 6.1: Resumo dos Métodos e Objetivos de Previsão no Policiamento Inteligente

Método	**Objetivo**	**Benefício**
Previsão de Curto Prazo	Desenvolvimento Tático	Melhora na resposta rápida a situações emergenciais
Previsão de Médio Prazo	Alocação de Recursos	Otimização da distribuição de recursos e pessoal
Previsão de Longo Prazo	Planejamento Estratégico	Planejamento e preparação para mudanças de longo prazo

O presente trabalho pode ser classificado como voltado para o curto e médio prazo, dado que visa tanto o desenvolvimento tático quanto a alocação de recursos, dependendo do tempo de consulta utilizado. A escolha deste horizonte temporal é motivada pela necessidade da polícia brasileira de uma ferramenta qualitativa que auxilie na tomada de decisões sobre estratégias táticas e alocação eficiente de recursos.

A sazonalidade criminal é um fenômeno importante que examina a variação da criminalidade em diferentes períodos, como meses, anos, feriados e eventos específicos (Cohen, 1941, Landau e Fridman, 1993, Ceccato, 2005, Hipp, 2003). A capacidade de prever crimes em períodos específicos é desejada e viável, considerando que o comportamento humano possui padrões previsíveis. Por exemplo, a incidência de roubos e furtos em áreas comerciais tende a aumentar devido ao grande fluxo de pessoas e ao volume de dinheiro

circulando. Da mesma forma, arrombamentos a casas de veraneio são mais comuns em períodos fora das férias ou do verão, quando essas propriedades estão frequentemente desocupadas. A identificação desses padrões sazonais é fundamental não apenas para prever crimes, mas também para gerenciar efetivamente a força policial, como a programação de treinamentos e férias durante períodos de menor criminalidade.

A análise de dados criminais é frequentemente usada para monitorar mudanças no comportamento criminal. As duas principais mudanças avaliadas são a alteração no nível e no padrão de criminalidade. A polícia frequentemente utiliza um período anual para comparar o nível criminal de um mês com o mesmo mês do ano anterior, calculando a diferença com a fórmula $\Delta_{12} = A_t - A_{t-12}$, onde A_t é a quantidade de crimes no mês atual e A_{t-12} no mesmo mês do ano passado (Armstrong, 1988, Deroeck, 1991, Mahmoud e Brown, 1992). A mudança no padrão criminal é analisada comparando o número de crimes previstos com os reais, usando a fórmula $P_{t\delta} = A_t - F_{t\delta}$, onde $F_{t\delta}$ representa o valor previsto.

No contexto da teoria das oportunidades para cometer crimes, métodos como a teoria da rotina de atividades, a teoria do padrão criminal e a perspectiva da escolha racional (Felson e Clarke, 1998, Cohen e Felson, 1979) são relevantes. A teoria da rotina de atividades (Cohen e Felson, 1979) sugere que as oportunidades para crimes estão concentradas em tempo e espaço, com três condições essenciais para a ocorrência de um crime: motivação dos criminosos, presença de alvos vulneráveis e ausência de um guardião eficaz. A previsibilidade do comportamento humano permite que crimes sejam previstos com base nesses fatores, o que é muito importante para a prevenção e a aplicação eficiente da lei.

Diversos estudos e pesquisas têm sido conduzidos na área de previsão de crimes (Camargo, 2008, Li, 2006, Mitchell et al., 2007), mostrando que, apesar das dificuldades associadas à variação na criminalidade, prever crimes oferece várias vantagens. A polícia pode planejar e executar ações táticas mais eficazes, como a determinação de áreas críticas e a alocação de recursos, enquanto o poder judiciário pode usar previsões para planejar novas unidades carcerárias e avaliar impactos das mudanças legislativas. A previsão de crimes, embora desafiadora e potencialmente custosa, proporciona benefícios significativos para a segurança pública e a gestão de recursos.

A presente dissertação abordou a integração de Redes Neurais Convolucionais (CNNs), Ontologia e Análise Preditiva com o intuito de aprimorar o policiamento inteligente. Ao longo deste estudo, formulamos e investigamos a hipótese de que a combinação sinérgica dessas tecnologias poderia ampliar significativamente a capacidade de detecção e prevenção de atividades cri-

minosas. Acreditamos que essa abordagem inovadora possui o potencial de otimizar tanto o planejamento estratégico quanto a roteirização das operações policiais, contribuindo assim para a construção de comunidades mais seguras.

6.1 Validação da Tese

Nossa tese sustentava que a implementação integrada de CNNs, Ontologia e Análise Preditiva tradicional oferece uma abordagem de inovação para o policiamento inteligente. Em particular, as CNNs podem processar informações visuais para identificar padrões complexos e comportamentos suspeitos, enquanto uma ontologia robusta aprimora a compreensão semântica dos dados, facilitando a interoperabilidade entre diferentes fontes de informação. A Análise Preditiva, por sua vez, emprega algoritmos avançados para não apenas reagir a eventos, mas também antecipar cenários futuros, permitindo um planejamento estratégico mais eficiente.

Os resultados obtidos corroboram essa tese, demonstrando que a integração dessas tecnologias proporciona uma gestão mais eficiente e proativa da segurança pública. A análise científica mostrou que a aplicação das CNNs no processamento de imagens pode efetivamente identificar padrões de comportamento criminoso mostrando-se muito promissor nesse cenário, enquanto a ontologia permite uma melhor integração e interpretação dos dados coletados. A Análise Preditiva tradicional se provou eficiente na antecipação de eventos para a alocação estratégica de recursos, validando a eficácia da nossa abordagem proposta.

6.2 Resposta à Hipótese

Confirmamos a hipótese de que a sinergia entre CNNs, Ontologia e Análise Preditiva pode melhorar sobremaneira a detecção e prevenção de atividades criminosas. A identificação eficaz de padrões, a representação semântica aprimorada e a capacidade preditiva avançada resultaram em um protótipo viável para o modelo de policiamento inteligente que não apenas responde a eventos, mas também antecipa e previne atividades criminosas com maior precisão.

6.3 Objetivos e Resultados

O objetivo geral desta dissertação foi propor um modelo integrado de policiamento inteligente que utiliza análise preditiva e redes neurais para melhorar a eficiência da Polícia Militar na prevenção e combate à criminalidade. Com base nos dados existentes, desenvolvemos o protótipo de um modelo que combina análise preditiva para prover a roteirização de patrulhas policiais de forma heurística, otimizando a tomada de decisão e a alocação de recursos.

Os objetivos específicos foram igualmente abordados e alcançados:

- **Desenvolvimento do Modelo:** Criamos o protótipo de um modelo de análise preditiva aprimorado pelas CNNs para identificar padrões de criminalidade, demonstrando sua eficácia na previsão de crimes em áreas geográficas específicas.
- **Integração Ontológica:** A integração de padrões ontológicos proporcionou uma visão mais completa e precisa, melhorando a capacidade dos gestores da Polícia Militar de tomar decisões informadas.
- **Aplicação Prática:** Avaliamos a aplicabilidade do método de análise preditiva para crimes passíveis de bonificação e investigamos fatores de risco e variáveis ambientais que afetam a criminalidade.
- **Ferramenta Integrada:** Propusemos uma ferramenta de previsão e otimização que pode ser integrada às atividades de policiamento preventivo, melhorando a eficiência operacional e a gestão de recursos.

6.4 Propostas para Trabalhos Futuros

Embora os resultados obtidos sejam promissores, há diversas áreas que podem ser exploradas em futuras pesquisas. Primeiramente, recomenda-se a ampliação do modelo para incluir dados de diferentes regiões e contextos, o que permitirá uma análise mais abrangente e adaptável às diversas realidades locais. Além disso, seria interessante investigar a integração de outras técnicas emergentes de machine learning, como redes neurais generativas ou algoritmos de aprendizado profundo, para aprimorar ainda mais a capacidade preditiva do sistema.

Outro aspecto relevante é a validação contínua da ferramenta em cenários reais e seu impacto na redução da criminalidade e na eficiência operacional. Estudos de caso adicionais e parcerias com forças policiais podem fornecer insights valiosos para ajustes e melhorias contínuas do modelo.

Finalmente, a implementação de um sistema de feedback que permita aos gestores da Polícia Militar ajustar o modelo com base em novas informações e alterações no padrão criminal pode aumentar a eficácia e a adaptabilidade da abordagem.

6.4.1 Validação Criminológica para Trabalhos Futuros

Para validar a eficácia do algoritmo de predição criminal, propõe-se um estudo de validação criminológica utilizando duas áreas com características sociocriminais semelhantes. Este método permitirá avaliar a precisão do algoritmo na previsão de crimes e a eficácia das recomendações de policiamento. O procedimento para esta validação será realizado conforme descrito a seguir:

6.4.1.1 Seleção das Áreas de Estudo

Selecionar duas áreas, denominadas Área A e Área B, que possuam perfis sociocriminais semelhantes. A escolha deve basear-se em dados históricos de crimes, características demográficas e outros fatores relevantes que garantam a comparabilidade entre as áreas.

6.4.1.2 Aplicação do Algoritmo

A aplicação do algoritmo será realizada da seguinte forma:

- **Área A:** Implementar o algoritmo de predição criminal e usar as previsões para definir a localização de máxima cobertura das viaturas policiais, com o objetivo de evitar a ocorrência de crimes. As viaturas serão posicionadas de acordo com as recomendações do algoritmo para maximizar a cobertura nas áreas de maior risco.
- **Área B:** Aplicar a mesma predição para prever onde os crimes podem ocorrer, mas sem realizar o policiamento predito. Ou seja, não serão alocadas viaturas de acordo com as recomendações do algoritmo. Esta área servirá como um controle para observar a eficácia da predição em um ambiente sem intervenção baseada nas recomendações do algoritmo.

6.4.1.3 Monitoramento e Avaliação

Após um período definido de monitoramento, os seguintes aspectos serão avaliados:

- **Área A:** Verificar se a alocação das viaturas predita pelo algoritmo resultou na redução da incidência de crimes. Se os crimes forem efetivamente evitados ou reduzidos, isso indicará que o policiamento foi eficiente e que a predição foi bem-sucedida.
- **Área B:** Avaliar a incidência de crimes na Área B. Se os crimes ocorrerem conforme previsto pelo algoritmo e a área não foi monitorada com base nas recomendações, isso sugerirá que a predição foi correta e que a ausência de policiamento recomendado contribuiu para a ocorrência dos crimes.

6.4.1.4 Inversão de Áreas

Para validar a tese de forma robusta, o estudo deve ser repetido com a inversão das áreas:

- **Área A** será usada como controle (sem policiamento predito) e **Área B** receberá a alocação de viaturas conforme as previsões do algoritmo.

- Após a aplicação do algoritmo nas áreas invertidas, as avaliações e análises serão realizadas novamente, seguindo o mesmo procedimento descrito anteriormente.

Essa abordagem permitirá uma análise comparativa da eficácia do algoritmo e das recomendações de policiamento em diferentes contextos, contribuindo para a validação de sua eficácia.

6.5 Considerações Finais

A integração de tecnologias emergentes, como Ontologia e Análise Preditiva, apromorada pelas CNN, representa sim um grande avanço no policiamento inteligente. A nossa pesquisa valida a hipótese de que essas tecnologias, quando combinadas de forma sinérgica, têm o potencial de melhorar a segurança pública. A abordagem proposta não só aprimora a capacidade de previsão e resposta a atividades criminosas, mas também contribui para uma gestão mais eficiente e estratégica da segurança. Esperamos que esta dissertação inspire novas pesquisas e inovações na área, promovendo comunidades mais seguras e resilientes.

Referências Bibliográficas

Plano de comando - versão 1 de 20dez17. 6a Seção do Estado-Maior da Polícia Militar do Estado de São Paulo, São Paulo, 2018a.

Plano diretor de tecnologia da informação e comunicação 2018 -2022. Diretoria de Telematica, São Paulo, 2018b.

ALLEN, P. *Mean square error: An analytical method*. Academic Press, 1971.

ALLEN-ZHU, Z.; LI, Y.; SONG, Z. A convergence theory for deep learning via over-parameterization. *arXiv preprint arXiv:1811.03962*, 2019.
Disponível em https://arxiv.org/abs/1811.03962

ARMSTRONG, J. S. Long-range forecasting: From crystal ball to computer. *Wiley*, 1988.

BAIG, M. A. A.; RATYAL, N. I.; AMIN, A.; JAMIL, U.; LIAQUAT, S.; KHALID, H. M.; ZIA, M. F. An ensemble deep cnn approach for power quality disturbance classification: A technological route towards smart cities using image-based transfer. *Future Internet*, v. 16, n. 12, 2024.
Disponível em https://www.mdpi.com/1999-5903/16/12/436

BALLESTEROS, P. R. Gestão de políticas de segurança pública no brasil: problemas, impasses e desafios. *Revista Brasileira de Segurança Pública*, v. 8, n. 1, p. 6–22, 2014.
Disponível em https://doi.org/10.15448/2177-6784.2015.2.22162

BASILIO, M. P.; PEREIRA, V.; COSTA, H. G. Método de apoio decisão multicritério: Um estudo empírico aplicado na classificação das Áreas integradas de segurança pública no estado do rio de janeiro. *Engevista*, v. 21, n. 1, p. 47–62, 2019.
Disponível em https://periodicos.uff.br/engevista/article/view/10129

BAYLEY, D. H. *Padrões de policiamento: Uma análise internacional comparativa*. São Paulo: Editora da Universidade de São Paulo, 2002.

BEATO FILHO, C. C. Ação e estratégia das organizações policiais. *Revista Unidade*, v. 1, n. 66, p. 79–100, 2009.

BISHOP, C. M. *Pattern recognition and machine learning*. Springer, 2006.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. Rio de Janeiro: LTC, 2000.

BREIMAN, L. Random forests. *Machine learning*, v. 45, n. 1, p. 5–32, 2001.

BRUHA, J.; FAMILI, A. Post-processing of data mining results. *Journal of Data Mining*, p. 123–135, 2000.

CALIS, J. O. G. Aplicação de redes neurais convolucionais para reconhecimento automático de placas de veículos. Trabalho de conclusão de curso (Bacharelado - Ciência da Computação), orientadora: Inês Aparecida Gasparotto Boaventura, 2018.

CAMARGO, E. E. A. Mapeamento do risco de homicídio com base na co-krigeagem binomial e simulação: um estudo de caso para são paulo, brasil. *Cadernos de Saúde Pública*, v. 24, n. 7, p. 1493–1508, 2008.

CAPLAN, J. M.; KENNEDY, L. W.; MILLER, J. Risk terrain modeling: Brokering criminological theory and gis methods for crime forecasting. *Justice Quarterly*, v. 28, n. 2, p. 360–381, 2011.

CARVALHO, M. D. S. M. V. D.; TONET, H. C. Qualidade na administração pública. *Revista de Administração Pública*, v. 28, n. 2, p. 137–152, 1994.

CECCATO, V. Homicide in sao paulo, brazil: Assessing spatial-temporal and weather variations. *Journal of Environmental Psychology*, v. 25, n. 3, p. 307–321, 2005.

CHAINEY, S.; TOMPSON, L.; UHLIG, S. The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal*, v. 21, n. 1–2, p. 4–28, 2008.
Disponível em https://link.springer.com/article/10.1057/palgrave.sj.8350066

COHEN, J. The geography of crime. *The Annals of the American Academy of Political and Social Science*, v. 217, n. 1, p. 29, 1941.

COHEN, L.; FELSON, M. Social change and crime rate trends: A routine activity approach. *American Sociological Review*, v. 44, p. 588–607, 1979.

COLDREN JR, J. R.; HUNTOON, A.; MEDARIS, M. Introducing smart policing: Foundations, principles, and practice. *Police Quarterly*, v. 16, n. 3, p. 275–286, 2013.

CONSTITUICAOSP Constituição do estado de são paulo. "Artigo 174 - Leis de iniciativa do Poder Executivo", 1989.

CORDNER, G. W.; GREENE, J. R.; BYNUM, T. S. Planejamento dos recursos humanos policiais. In: *Administração do Trabalho Policial: Questões e Análises*, Polícia, Segurança Pública, p. 61–83, 2007.

COUTO, G. Mapas de ativação empilhados para próxima camada. *Revista de Inteligência Artificial Aplicada*, v. 1, n. 1, p. 1–10, 2023.
Disponível em https://exemplo.com/mapas-de-ativacao

DEROECK, R. Is there a gap between forecasting theory and practice? a personal view. *International Journal of Forecasting*, v. 6, p. 17–19, 1991.

DEWINTER, M.; VANDEVIVER, C.; BEKEN, T. V.; WITLOX, F. Analysing the police patrol routing problem: A review. *ISPRS International Journal of Geo-Information*, v. 9, n. 3, p. 157, 2020.

FELSON, M.; CLARKE, R. *Opportunity makes the thief.* Home Office, Police Research Group, 1998.

FREITAS, J. Direito fundamental à boa administração pública e o reexame dos institutos da autorização de serviço público, da convalidação e do poder de polícia administrativa. In: ARAGÃO, A. S. D.; MARQUES NETO, F. D. A., eds. *Direito Administrativo e seus novos paradigmas*, Belo Horizonte: Fórum, p. 311–334, 2008.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning.* MIT Press, 2016.
Disponível em http://www.deeplearningbook.org/

GRABOSKY, P. N. Fear of crime, and fear reduction strategies. *Current issues in criminal justice*, v. 7, n. 1, p. 7–19, 1995.
Disponível em https://www.aic.gov.au/sites/default/files/2020-05/tandi044.pdf

GRUBER, T. R. A translation approach to portable ontology specifications. *Knowledge Acquisition*, v. 5, n. 2, p. 199–220, 1993.

GUARINO, N. Formal ontology in information systems. *Proceedings of FOISi98, Trento, Italy, 6-8.*, v. I, p. 3–15, June 1998.

GUSMÃO, A. P. H. D.; COSTA BORBA, B. F. D.; CLEMENTE, T. R. N. Management information system for police facility location. In: *International Conference on Decision Support System Technology*, Springer, Cham, 2020, p. 86–98.

HART, T.; ZANDBERGEN, P. Kernel density estimation and hotspot mapping: Examining the influence of interpolation method, grid cell size, and bandwidth on crime forecasting. *Policing: An International Journal of Police Strategies & Management*, 2014.

HAYKIN, S. *Redes neurais: princípios e prática*. Porto Alegre: Bookman, 2001.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 770–778, 2016.
Disponível em https://ieeexplore.ieee.org/document/7780459

HIPP, J. E. A. Crimes of opportunity or crimes of emotion-testing two explanations of seasonal change in crime. *Social Forces*, v. 82, p. 1333–1372, 2003.

ISHAK, P. W. Murder nature: Weather and violent crime in rural brazil. *World Development*, 2022.

JAIN, A. K. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, v. 31, p. 651–666, 2010.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. *Science*, v. 349, n. 6245, p. 255–260, 2015.

KASABOV, N. *Foundations of neural networks, fuzzy systems, and knowledge engineering*. The MIT Press, 1996.

KENNEDY, L. W.; CAPLAN, J. M.; PIZA, E. Risk clusters, hotspots, and spatial intelligence: risk terrain modeling as an algorithm for police resource allocation strategies. *Journal of Quantitative Criminology*, v. 27, n. 3, p. 339–362, 2011.

KOPER, C. S. Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice quarterly*, v. 12, n. 4, p. 649–672, 1995.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Learning multiple layers of features from tiny images. *Technical Report*, cIFAR-10 dataset, 2009.
Disponível em https://www.cs.toronto.edu/~kriz/cifar.html

LANDAU, S.; FRIDMAN, D. The seasonality of violent crime: The case of robbery and homicide in israel. *Journal of Research in Crime and Delinquency*, v. 30, n. 2, p. 163, 1993.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, v. 86, n. 11, p. 2278–2324, 1998.
Disponível em https://ieeexplore.ieee.org/document/726791

LEI4320 Lei nž 4.320, de 17 de março de 1964. Estatui as normas gerais de direito financeiro para elaboração e controle dos orçamentos e balanços da União, dos Estados, dos Municípios e do Distrito Federal, estabelece a Lei de Orçamento como instrumento para a discriminação da receita e despesa, conforme observa-se no teor de seu artigo 2ž., 1964.
Disponível em http://www.planalto.gov.br/ccivil_03/leis/l4320.htm

LI, S.-T. E. A. A knowledge discovery approach to supporting crime prevention. In: *JCIS-2006 Proceedings*, s.n., 2006.

LOA Lei orçamentária anual (loa). Orça a receita e fixa a despesa do Estado para o exercício anual, compreendendo, nos termos do artigo 174, ğ 4ž, da Constituição Estadual., 2023.

MAHMOUD, A.; BROWN, M. Techniques of forecasting. *International Journal of Forecasting*, 1992.

MEIJER, A.; WESSELS, M. Predictive policing: Review of benefits and drawbacks. *International Journal of Public Administration*, v. 42, n. 12, p. 1031–1039, 2019.
Disponível em http://10.1080/01900692.2019.1575664

MICHALSKI, R. S.; CARBONELL, J. G.; MITCHELL, T. M. *Machine learning: An artificial intelligence approach*. [S.l.]: Morgan Kaufmann Publishers, 1986.

MITCHELL, M.; BROWN, D.; CONKLIN, J. A crime forecasting tool for the web-based crime analysis toolkit. In: *IEEE Systems and Information Engineering Design Symposium, 2007. SIEDS 2007*, s.n., 2007, p. 1–5.

MOSES, L. B.; CHAN, J. Algorithmic prediction in policing: assumptions, evaluation, and accountability. *Policing and Society*, v. 28, n. 7, p. 806–822, 2018.

OLIVEIRA, W. A. D.; MORETTI, A. C.; REIS, E. F. Multi-vehicle covering tour problem: building routes for urban patrolling. *Pesquisa Operacional*, v. 35, n. 3, p. 617–644, 2015.

DE OLIVEIRA CAIRES, D. *Técnicas de interpretabilidade para aprendizado de máquina: um estudo abordando avaliação de crédito e detecção de fraude*. Dissertação de Mestrado, acesso em: 16 de dezembro de 2022, 2022. Disponível em https://www.teses.usp.br/teses/disponiveis/55/55137/tde-16122022-180337/pt-br.php

OLLIGSCHLAEGER, A. *Neural networks: A comprehensive foundation*. IEEE Press, 1997.

PEARSON, K. Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, v. 58, p. 240–242, 1895.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, É. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.

PERRY, W. L. *Predictive policing: The role of crime forecasting in law enforcement operations*. Relatório Técnico, Rand Corporation, 2013. Disponível em https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf

PITERI, M. A.; RODRIGUES, J. C. *Fundamentos de visão computacional*. Presidente Prudente: FCT/UNESP-PP, 2011.

PMESP2022 Sistema de gestão da polícia militar do estado de são paulo (gespol). Acesso em dezembro de 2023, 2022. Disponível em http://www.intranet.polmil.sp.gov.br/organizacao/unidades/6empm/documentos/Documentos/livrogespol.pdf

PYTHON SOFTWARE FOUNDATION Python. Acesso em: 2021, 2021. Disponível em https://www.python.org/

RATCLIFFE, J. Crime mapping: Spatial and temporal challenges. In: *Handbook of Quantitative Criminology*, Springer, New York, NY, 2010. Disponível em https://link.springer.com/chapter/10.1007/978-0-387-77650-7_2

RIEDMILLER, M.; BRAUN, H. A direct adaptive method for faster backpropagation learning: The rprop algorithm. *IEEE International Conference on Neural Networks*, p. 586–591, 1993.

RODRIGUES, F. A.; MACIEL, C. Um método para captura e compartilhamento de dados abertos educacionais via um processo etl. *Anais do CSBC 2022.*, 2022.

RUMMENS, A.; HARDYNS, W.; PAUWELS, L. The use of predictive analysis in spatiotemporal crime forecasting: Building and testing a model in an urban context. *Applied Geography*, v. 86, p. 255–261, 2017.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. Disponível em https://arxiv.org/abs/1409.1556

SIMPSON, N. C.; HANCOCK, P. G. Fifty years of operational research and emergency response. *Journal of the Operational Research Society*, v. 60, n. sup1, p. S126–S139, 2009.

SINCLAIR, C.; DAS, S. Traffic accidents analytics in uk urban areas using k-means clustering for geospatial mapping. In: *2021 International Conference on Sustainable Energy and Future Electric Transportation (SEFET)*, IEEE, 2021, p. 1–7.

SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; PARIKH, D.; RUSSAKOVSKY, O.; ET AL. Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 1–9, 2015. Disponível em https://ieeexplore.ieee.org/document/7163122

U.M *Advances in knowledge discovery and data mining.* AI Magazine: From Data Mining to Knowledge, 1996.

USCHOLD, M.; GRUNINGER, M. Ontologies: Principles, methods and applications. *Knowledge Engineering Review*, 1996.

VALASIK, M.; BRAULT, E. E.; MARTINEZ, S. M. Forecasting homicide in the red stick: Risk terrain modeling and the spatial influence of urban blight on lethal violence in baton rouge, louisiana. *Journal of Urban Affairs*, v. 43, n. 2, p. 195–212, 2021.

WIDROW, B.; LEHR, M. A. *1990s: The delta rule.* IEEE Press, 1990.

WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, v. 30, n. 1, p. 79–82, 2005.

WINKLHOFER, H.; DIAMANTOPOULOS, A.; WITT, J. Framework for the application of forecasting techniques. *European Journal of Operational Research*, 1996.

APÊNDICE

A

Procedimentos de Tratamento de Dados

Este apêndice apresenta, em detalhes, os procedimentos de pré-processamento e tratamento dos dados utilizados no desenvolvimento deste trabalho. São descritas as etapas de importação, transformação, filtragem, codificação e visualização dos dados, com os respectivos trechos de código implementados em Python.

A.1 Tratamento dos Dados

Os dados utilizados neste estudo foram coletados da Secretaria da Segurança Pública e da PMESP. Eles incluem informações sobre a localização, tipo e data dos crimes ocorridos em São Paulo. A base de dados foi limpa e organizada para facilitar a análise.

A.1.1 Pré-processamento

O pré-processamento dos dados é uma etapa crucial para garantir que o modelo de machine learning receba dados de qualidade. O código abaixo realiza uma série de operações de pré-processamento em um conjunto de dados de criminalidade utilizando bibliotecas populares de Python como pandas e scikit-learn. O objetivo dessas operações é preparar os dados para uma análise ou um modelo de aprendizado de máquina. A seguir, descrevemos cada passo do código:

```
import pandas as pd
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.model_selection import train_test_split
import holidays
```

```
import warnings

warnings.filterwarnings('ignore')
```

O código inicia importando diversas bibliotecas essenciais:

- **pandas** para manipulação de dados.
- **StandardScaler** e **LabelEncoder** do `sklearn.preprocessing` para normalização e codificação dos dados.
- **train_test_split** do `sklearn.model_selection` para dividir o conjunto de dados em treino e teste.
- **holidays** para a manipulação de feriados.
- **warnings** para ignorar avisos que podem surgir durante a execução do código.

O código suprime todos os avisos utilizando `warnings.filterwarnings('ignore')`, o que ajuda a manter o foco nos resultados sem distrações causadas por mensagens de aviso.

A.1.2 Carregamento

Os dados são carregados a partir de um arquivo CSV chamado `crime_data.csv` para um DataFrame do `pandas` utilizando o comando:

```
df = pd.read_csv('crime_data.csv')
```

A.1.3 Renomear Colunas

As colunas do DataFrame são renomeadas para facilitar a manipulação. O código a seguir demonstra como renomear as colunas do DataFrame `df`:

```
# Renomear as colunas do DataFrame
df = df.rename(columns={
    'CIDADE': 'cidade',
    'DATA_OCORRENCIA_BO': 'data',
    'HORA_OCORRENCIA_BO': 'hora',
    'LATITUDE': 'latitude',
    'LONGITUDE': 'longitude',
    'NATUREZA_APURADA': 'descricao'
})
```

A.1.4 Função para Extrair Hora

Uma função é definida para extrair a hora de uma string:

```
def get_hora(input) -> int:
    try:
        return int(input[0:2])
    except Exception as e:
        return -1
```

A.1.5 Filtrar Dados Relevantes

Os dados são filtrados para incluir apenas certas classes de crimes e apenas registros da cidade de São Paulo. O código a seguir demonstra como filtrar os dados para essas condições:

```
# Definir as classes de crimes de interesse
classes_crimes = [
    'FURTO - OUTROS',
    'FURTO DE CARGA',
    'FURTO DE VEÍCULO',
    'ROUBO - OUTROS',
    'ROUBO DE CARGA',
    'ROUBO DE VEÍCULO'
]

# Filtrar o DataFrame para incluir apenas as classes de crimes
   especificadas
df = df.loc[df['descricao'].isin(classes_crimes)]

# Filtrar o DataFrame para incluir apenas os registros da cidade de São
   Paulo
df = df.loc[df['cidade'] == 'S.PAULO']
```

A.1.6 Processar Colunas de Data e Hora

As colunas de data e hora são processadas para extração de informações adicionais:

```
df['data'] = pd.to_datetime(df['data'])
df['hora'] = df['hora'].astype(str).apply(get_hora)
df['dia'] = df['data'].dt.day
df['mes'] = df['data'].dt.month
df['ano'] = df['data'].dt.year
df['dia_semana'] = df['data'].dt.weekday
```

A.1.7 Codificar a Coluna `descricao`

A coluna `descricao` é codificada utilizando `LabelEncoder`:

```
le = LabelEncoder()
df['descricao'] = le.fit_transform(df['descricao'])
```

A.1.8 One-Hot Encoding para `dia_semana`

A coluna `dia_semana` é transformada utilizando `One-Hot Encoding`:

```
df = pd.get_dummies(df, columns=['dia_semana']).reset_index()
```

A.1.9 Adicionar Feriados

Os feriados são adicionados ao DataFrame para enriquecer os dados com informações sobre feriados nacionais. O código a seguir demonstra como adicionar esses feriados:

```
# Obter feriados nacionais do Brasil
feriados = holidays.Brazil()

# Selecionar feriados para o período de 2023 a 2024
feriados2023 = feriados['2023-01-01': '2024-12-31']

# Converter os feriados para um DataFrame
df_feriados = pd.DataFrame(data=feriados2023, columns=['data'])

# Converter a coluna de datas para o formato datetime
df_feriados['data'] = pd.to_datetime(df_feriados['data'])

# Adicionar uma coluna indicando se o dia é um feriado
df_feriados['e_feriado'] = 1

# Mesclar o DataFrame original com o DataFrame de feriados
df = pd.merge(df, df_feriados, on='data', how='left').reset_index()

# Substituir valores nulos na coluna 'e_feriado' por 0
df['e_feriado'] = df['e_feriado'].fillna(0)
```

Após a execução do código, os seguintes feriados foram adicionados ao DataFrame:

```
[datetime.date(2023, 1, 1), datetime.date(2023, 4, 7), datetime.date(2023,
    4, 21),
 datetime.date(2023, 5, 1), datetime.date(2023, 9, 7), datetime.date(2023,
     10, 12),
 datetime.date(2023, 11, 2), datetime.date(2023, 11, 15), datetime.date
     (2023, 12, 25),
 datetime.date(2024, 1, 1), datetime.date(2024, 3, 29), datetime.date(2024,
     4, 21),
 datetime.date(2024, 5, 1), datetime.date(2024, 9, 7), datetime.date(2024,
     10, 12),
 datetime.date(2024, 11, 2), datetime.date(2024, 11, 15), datetime.date
     (2024, 12, 25)]
```

E o DataFrame resultante contém as seguintes colunas:

```
Index(['level_0', 'index', 'cidade', 'data', 'hora', 'latitude', 'longitude
    ',
       'descricao', 'dia', 'mes', 'ano', 'dia_semana_0.0', 'dia_semana_1
           .0',
       'dia_semana_2.0', 'dia_semana_3.0', 'dia_semana_4.0', 'dia_semana_5
           .0',
       'dia_semana_6.0', 'e_feriado'],
      dtype='object')
```

A.1.10 Criar Variáveis de Atraso para Feriados

Variáveis de atraso para feriados são criadas para avaliar o impacto dos feriados em diferentes períodos anteriores. O código a seguir cria variáveis que indicam se o dia é um feriado, considerando os feriados dos últimos 3 dias e dos próximos 3 dias:

```
# Inicializar lista para armazenar os nomes das variáveis de atraso
delay_features = []

# Criar variáveis de atraso para feriados
for i in range(-3, 4):
    df['e_feriado' + str(i)] = df['e_feriado'].shift(i)
    delay_features.append('e_feriado' + str(i))
```

Com a execução deste código, variáveis de atraso são adicionadas ao DataFrame para capturar o efeito dos feriados em dias anteriores e posteriores, permitindo uma análise mais detalhada do impacto dos feriados no conjunto de dados.

A.1.11 Selecionar Features para o Modelo

As features são selecionadas para o modelo com o objetivo de incluir todas as variáveis relevantes que serão utilizadas durante o treinamento. O código a seguir mostra como as features são definidas e inclui uma lista das variáveis que serão utilizadas:

```
# Definir as features a serem usadas no modelo
features = [
    'hora',
    'latitude',
    'longitude',
    'dia',
    'mes',
    'ano',
    'dia_semana_0.0',
    'dia_semana_1.0',
    'dia_semana_2.0',
    'dia_semana_3.0',
```

```
    'dia_semana_4.0',
    'dia_semana_5.0',
    'dia_semana_6.0',
    'e_feriado'
] + delay_features
```

Após a execução do código, as features selecionadas para o modelo são:

```
['hora', 'latitude', 'longitude', 'dia', 'mes', 'ano', 'dia_semana_0.0',
 'dia_semana_1.0', 'dia_semana_2.0', 'dia_semana_3.0', 'dia_semana_4.0',
 'dia_semana_5.0', 'dia_semana_6.0', 'e_feriado', 'e_feriado-3', 'e_feriado
    -2',
 'e_feriado-1', 'e_feriado0', 'e_feriado1', 'e_feriado2', 'e_feriado3']
```

Essas features incluem variáveis temporais e geoespaciais, bem como indicadores de feriados e suas versões deslocadas, que são importantes para capturar padrões e influências temporais nos dados.

A.1.12 Excluir Valores Ausentes

Valores ausentes são excluídos:

```
df = df.dropna()
```

A.1.13 Visualização da Correlação entre Variáveis

O código a seguir ilustra como calcular e visualizar a correlação entre um conjunto específico de variáveis em um DataFrame utilizando Python. A visualização é feita por meio de um mapa de calor, que facilita a interpretação das relações entre as variáveis selecionadas:

```
# Importar as bibliotecas necessárias para visualização
import seaborn as sns
import matplotlib.pyplot as plt

# Calcular e visualizar a correlação entre as variáveis selecionadas
# A função .corr() calcula a matriz de correlação entre as variáveis
   numéricas
sns.heatmap(
    df[['hora', 'latitude', 'longitude', 'descricao', 'dia', 'mes', 'ano',
        'e_feriado', 'e_feriado-3', 'e_feriado-2', 'e_feriado-1',
        'e_feriado1', 'e_feriado2', 'e_feriado3']].corr(),
    annot=True,          # Adicionar valores numéricos no mapa de calor
    fmt='.1f',           # Formatar os valores como números decimais com uma
        casa decimal
    cmap='Blues'         # Definir a paleta de cores para o mapa de calor
)

# Adicionar um título ao gráfico
plt.title('Correlação das Variáveis da Base de Dados')
```

```
# Exibir o gráfico
plt.show()
```

A.1.13.1 Correlação de Pearson

A correlação de Pearson é uma medida estatística que avalia a força e a direção de uma relação linear entre duas variáveis quantitativas. Este coeficiente varia de -1 a 1, onde valores próximos a 1 indicam uma forte correlação positiva, valores próximos a -1 indicam uma forte correlação negativa, e valores próximos a 0 sugerem pouca ou nenhuma correlação linear entre as variáveis analisadas.

- **Cálculo da Correlação:** A função a seguir calcula a matriz de correlação de Pearson entre as variáveis selecionadas no DataFrame `df`. A matriz de correlação mede a força e a direção das relações lineares entre pares de variáveis.

```
df[['hora', 'latitude', 'longitude', 'descricao',
'dia', 'mes', 'ano', 'e_feriado', 'e_feriado-3',
'e_feriado-2', 'e_feriado-1', 'e_feriado1',
'e_feriado2', 'e_feriado3']].corr()
```

Na FiguraA.1, as variáveis estão dispostas em uma matriz onde os coeficientes de correlação são representados por uma escala de cores. Cores mais escuras indicam correlações mais fortes, tanto positivas quanto negativas. Este tipo de visualização é fundamental para identificar padrões e relações entre variáveis, o que pode fornecer insights valiosos para a modelagem preditiva e análise de dados.

Além disso, o gráfico permite uma rápida interpretação visual das interações entre variáveis, facilitando a detecção de possíveis redundâncias e colinearidades, que são aspectos críticos na construção de modelos estatísticos e de aprendizado de máquina. A análise das correlações apresentadas pode orientar a seleção de variáveis para modelos preditivos, influenciar decisões sobre o tratamento de dados e contribuir para uma compreensão mais aprofundada das dinâmicas subjacentes ao fenômeno estudado.

A.2 Preparação dos Dados

A.2.1 Separar Dados em Treino e Teste

```
# Importar a função para dividir os dados
from sklearn.model_selection import train_test_split

# Dividir o DataFrame em conjuntos de treino e teste
# 'df[features]' contém as características de entrada
```

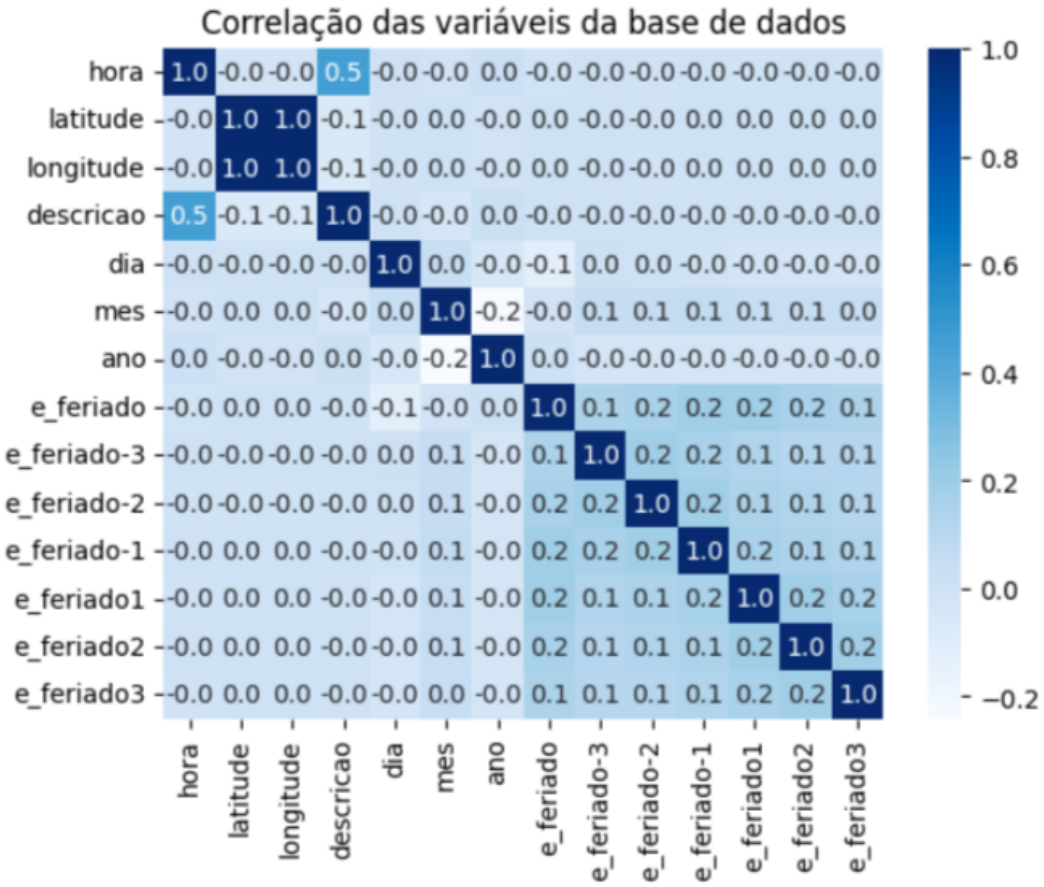


Figura A.1: Correlação de Pearson entre variáveis.

```
# 'df['descricao']' contém os rótulos (classes) para a previsão
# 'test_size=0.3' indica que 30% dos dados serão usados para teste
train_x, test_x, train_y, test_y = train_test_split(df[features], df['
    descricao'], test_size=0.3)
```

A.2.2 Escalonamento dos Dados

```
# Importar o escalonador padrão
from sklearn.preprocessing import StandardScaler

# Instanciar o escalonador
scaler = StandardScaler()

# Ajustar o escalonador aos dados de treino e transformar os dados de
    treino
# 'fit_transform' ajusta o escalonador aos dados e aplica a transformação
train_x = scaler.fit_transform(train_x)

# Transformar os dados de teste usando o mesmo escalonador
# 'transform' aplica a mesma transformação aos dados de teste
test_x = scaler.transform(test_x)
```

A.2.3 Análise Espacial

Utilizamos a biblioteca `folium` para visualização geoespacial e `geopandas` para manipulação de dados geoespaciais.

A.2.4 Carregar Dados Geoespaciais

```
# Importar bibliotecas necessárias para análise geoespacial
import folium
from folium.plugins import HeatMap
import geopandas as gpd

# Suponhamos que 'df' seja um DataFrame com dados de localização
# Criar um GeoDataFrame a partir do DataFrame df
# 'longitude' e 'latitude' são colunas no DataFrame df contendo coordenadas
    geográficas
# 'geometry' define a geometria dos pontos baseados nas coordenadas
gdf = gpd.GeoDataFrame(df, geometry=gpd.points_from_xy(df.longitude, df.
    latitude))
```

A.2.5 Mapa de Calor

```
# Criar um mapa centrado em uma coordenada genérica com um nível de zoom
    inicial
# As coordenadas fornecidas são genéricas e devem ser ajustadas conforme a
    localização desejada
mapa = folium.Map(location=[0.00000, 0.00000], zoom_start=12)

# Preparar os dados para o mapa de calor
# Extrair as coordenadas de latitude e longitude dos pontos no GeoDataFrame
# 'heat_data' é uma lista de listas contendo as coordenadas (latitude,
    longitude)
heat_data = [[point.y, point.x] for point in gdf.geometry]

# Adicionar o mapa de calor ao mapa principal
# 'HeatMap' cria uma camada de mapa de calor usando os dados de coordenadas
HeatMap(heat_data).add_to(mapa)

# Salvar o mapa como um arquivo HTML
# 'heatmap.html' é o nome do arquivo onde o mapa será salvo
mapa.save('heatmap.html')
```

A.3 Modelagem com Machine Learning

Para a modelagem, foram utilizados diversos algoritmos de classificação para prever o tipo de crime. A seguir, descrevemos o processo de treinamento e avaliação dos modelos, bem como a análise dos resultados obtidos.

A.3.1 Importação e Definição dos Classificadores

Inicialmente, importamos as bibliotecas necessárias e definimos os classificadores a serem utilizados no experimento. O código a seguir configura os classificadores e prepara o ambiente para o treinamento e avaliação:

```
# Importar as bibliotecas necessárias para avaliação e modelagem
from sklearn.metrics import confusion_matrix, balanced_accuracy_score,
    accuracy_score
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.ensemble import RandomForestClassifier

# Lista de classificadores
classifiers = {
    'DecisionTreeClassifier()': DecisionTreeClassifier(),
    'LogisticRegression()': LogisticRegression(),
    'KNeighborsClassifier()': KNeighborsClassifier(),
    'RandomForestClassifier()': RandomForestClassifier(),
    'GradientBoostingClassifier()': GradientBoostingClassifier(),
    'GaussianNB()': GaussianNB(),
}
```

A.3.2 Treinar Modelos

O código a seguir realiza o pré-processamento dos dados, que inclui a limpeza, a divisão dos dados e o escalonamento das variáveis. Primeiramente, a linha `df = df.dropna()` remove as linhas com valores ausentes do DataFrame, garantindo que o conjunto de dados esteja completo e livre de lacunas antes do treinamento do modelo. Em seguida, a instrução `(train_x, test_x, train_y, test_y) = train_test_split(df[features], df['descricao'], test_size=0.3)` divide o DataFrame em conjuntos de treinamento e teste, alocando 30% dos dados para o teste e 70% para o treinamento. Esta divisão é essencial para avaliar o desempenho do modelo em dados que não foram usados durante o treinamento. Para garantir que todas as variáveis preditoras estejam na mesma escala, utiliza-se o `StandardScaler`. A linha `scaler = StandardScaler()` cria uma instância do escalonador, `train_x = scaler.fit_transform(train_x)` ajusta o escalonador aos dados de treinamento e transforma esses dados, enquanto `test_x = scaler.transform(test_x)` aplica a mesma transformação aos dados de teste. O escalonamento é importante para melhorar a performance dos algoritmos de Aprendizado de Máquina e assegurar que todas as variáveis sejam comparadas de forma justa.

```
# Treinar e avaliar modelos
for name, clf in classifiers.items():
    clf.fit(train_x, train_y)
    pred_y = clf.predict(test_x)
```

```
cm = confusion_matrix(test_y, pred_y)
bal_acc = balanced_accuracy_score(test_y, pred_y)
acc = accuracy_score(test_y, pred_y)
```

A.4 Avaliação do Modelo

Para avaliar o desempenho dos modelos de machine learning, utilizamos várias métricas cruciais que fornecem uma visão abrangente sobre a eficácia dos classificadores. A matriz de confusão foi empregada para ilustrar o número de previsões corretas e incorretas em cada classe, permitindo uma análise detalhada das classificações verdadeiras e falsas. A acurácia balanceada foi calculada para medir a precisão do modelo considerando o desequilíbrio entre as classes, fornecendo uma média das acurácias de cada classe e garantindo uma avaliação justa mesmo em cenários com classes desbalanceadas. A acurácia geral foi também reportada para quantificar a proporção de previsões corretas em relação ao total, oferecendo uma visão geral da performance do modelo. Adicionalmente, a importância das features foi analisada, especialmente com o `RandomForestClassifier`, para identificar quais variáveis contribuíram mais para as previsões do modelo, auxiliando na compreensão de quais fatores são mais relevantes para a previsão dos crimes e permitindo a potencial otimização do modelo ao focar nas variáveis mais impactantes.

Os resultados obtidos para cada classificador são os seguintes:

A.4.1 Avaliação dos Modelos

Os resultados dos modelos de classificação foram avaliados com base na matriz de confusão, acurácia balanceada e acurácia geral. A seguir, apresentamos os resultados obtidos para cada classificador.

- **Heatmap da Matriz de Confusão**:

```
# Heatmap da matriz de confusão
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues')
plt.title(f'Matriz de Confusão - {name}')
plt.xlabel('Previsto')
plt.ylabel('Real')
plt.show()
```

- **DecisionTreeClassifier()**:

```
CM DecisionTreeClassifier() Original:
[[21703    11  2082  6225   145   658]
 [   15     0     3    11     0     1]
 [ 2084     1  1515  1772    73   296]
 [ 6133     6  1724  8851   190  1000]
 [  149     0    50   156    43    27]
```

```
 [  641     0   305   928    27   249]]
BalAcc DecisionTreeClassifier() Original: 0.2798891245025692
Acc DecisionTreeClassifier() Original: 0.5670007358867435
```

O `DecisionTreeClassifier` apresentou uma acurácia de 56.7% e uma acurácia balanceada de 27.9%. A matriz de confusão indica que o modelo teve dificuldades em classificar corretamente algumas categorias, particularmente para classes menos representadas.

- **LogisticRegression()**:

```
CM LogisticRegression() Original:
[[24628     0     4  6192     0     0]
 [   26     0     0     4     0     0]
 [ 3116     0     5  2620     0     0]
 [ 8506     0     3  9395     0     0]
 [  334     0     0    91     0     0]
 [  898     0     2  1250     0     0]]
BalAcc LogisticRegression() Original: 0.22076696738266746
Acc LogisticRegression() Original: 0.5962084311595472
```

O `LogisticRegression` alcançou uma acurácia de 59.6% e uma acurácia balanceada de 22.1%. Embora a acurácia geral seja razoável, a baixa acurácia balanceada sugere que o modelo pode estar tendendo a classes mais frequentes.

- **KNeighborsClassifier()**:

```
CM KNeighborsClassifier() Original:
[[24848     0  1135  4795     3    43]
 [   18     0     4     8     0     0]
 [ 3265     0   720  1726     2    28]
 [ 8468     0  1191  8163     3    79]
 [  268     0    40   113     3     1]
 [  970     0   199   966     0    15]]
BalAcc KNeighborsClassifier() Original: 0.23358433190425945
Acc KNeighborsClassifier() Original: 0.591320040648982
```

O `KNeighborsClassifier` apresentou uma acurácia de 59.1% e uma acurácia balanceada de 23.4%. Similar ao `LogisticRegression`, o modelo mostrou boa acurácia geral, mas teve desempenho inferior nas classes menos frequentes.

- **RandomForestClassifier()**:

```
CM RandomForestClassifier() Original:
[[24452     1   847  5413    17    94]
 [   17     0     2    11     0     0]
 [ 2516     0  1111  2058     3    53]
```

```
 [ 6053     0   657 11029    12   153]
 [  192     0    29   194     5     5]
 [  635     0   126  1336     3    50]]
BalAcc RandomForestClassifier() Original: 0.27297106229293494
Acc RandomForestClassifier() Original: 0.6420962259522724
```

O `RandomForestClassifier` teve a melhor acurácia geral de 64.2% e uma acurácia balanceada de 27.3%. Isso demonstra que o modelo tem uma boa capacidade de generalização e é relativamente equilibrado na classificação das diferentes classes.

- **GradientBoostingClassifier()**:

```
CM GradientBoostingClassifier() Original:
[[24721   206   409  5471    12     5]
 [   20     2     1     7     0     0]
 [ 2811    50   742  2132     3     3]
 [ 6173   207   390 11103    19    12]
 [  194     7    19   198     5     2]
 [  556    48    78  1455     4     9]]
BalAcc GradientBoostingClassifier() Original: 0.27233481282508765
Acc GradientBoostingClassifier() Original: 0.6409573536111014
```

O `GradientBoostingClassifier` apresentou uma acurácia de 64.1% e uma acurácia balanceada de 27.2%. Assim como o `RandomForestClassifier`, este modelo mostrou um desempenho sólido, com boa acurácia geral.

- **GaussianNB()**:

```
CM GaussianNB() Original:
[[ 4528 22565   869  2862     0     0]
 [    0    29     1     0     0     0]
 [  234  4703   200   604     0     0]
 [  912 14011   204  2777     0     0]
 [    5   402     5    13     0     0]
 [   95  1543    64   448     0     0]]
BalAcc GaussianNB() Original: 0.2172512213592673
Acc GaussianNB() Original: 0.13200406489820232
```

O `GaussianNB` obteve a menor acurácia geral de 13.2% e uma acurácia balanceada de 21.7%. A baixa performance do `GaussianNB` sugere que a suposição de independência entre as variáveis pode não ser válida para este problema.

A.4.2 Importância das Features

Além da avaliação dos classificadores, a importância das features foi analisada utilizando o `RandomForestClassifier`. A seguir está o código utilizado para calcular e plotar a importância das features:

```
# Importância das features
clf_rf = classifiers['RandomForestClassifier()']
feature_importance = clf_rf.feature_importances_
features_df = pd.DataFrame({'name': features, 'importance':
    feature_importance})
features_df = features_df.sort_values(by='importance', ascending=False)

# Plotar importância das features
plt.figure(figsize=(10, 6))
plt.barh(features_df['name'], features_df['importance'])
plt.xlabel('Importância')
plt.title('Importância das Features')
plt.gca().invert_yaxis()
plt.show()
```

A importância das *features* fornece *insights* sobre quais variáveis têm mais influência na previsão dos crimes. O gráfico resultante permite identificar as features mais relevantes, o que pode ajudar a melhorar a modelagem ao focar nas variáveis mais significativas.

A.5 Geração Via Análise Preditiva de Mapas com Grid

No software de policiamento preditivo que será desenvolvido a partir desta pesquisa, a criação de um mapa será a maneira padrão e mais importante de representar o risco de crime. O mapeamento de crimes tem uma longa história no trabalho policial, e software de policiamento preditivo pretende fazer uma representação da distribuição espacial do crime e aprimorar sua visualização com as capacidades das tecnologias digitais para tornar visível a atividade criminosa em toda o setor de estudo e, assim, identificar padrões e/ou clusters de crime:

É desnecessário dizer que os alfinetes não são mais colocados nos mapas. Os padrões serão identificados pelo software de policiamento preditivo diretamente nos dados, e as possibilidades que as tecnologias digitais proporcionam para representar visualmente esses padrões.

A Figura acima fornece alguns exemplos de como o risco pode ser colocado em um mapa (por motivo de confidencialidade do trabalho da Polícia Paulista, o local de geração foi plotado de forma aleatória para a não identificação do local de estudo). Nenhum deles é dado ou definido. Todos usam técnicas diferentes para destacar aspectos específicos da conexão presumida entre risco e espaço fortemente voltado para a representação visual do risco de crime na forma de um mapa. Esse mapa com geração de Grid pode ser plotado preditivamente dentro de um setor predeterminado ond são analisadas todas as possibilidades de localização de patrulhas policiais para se evitar o crime.

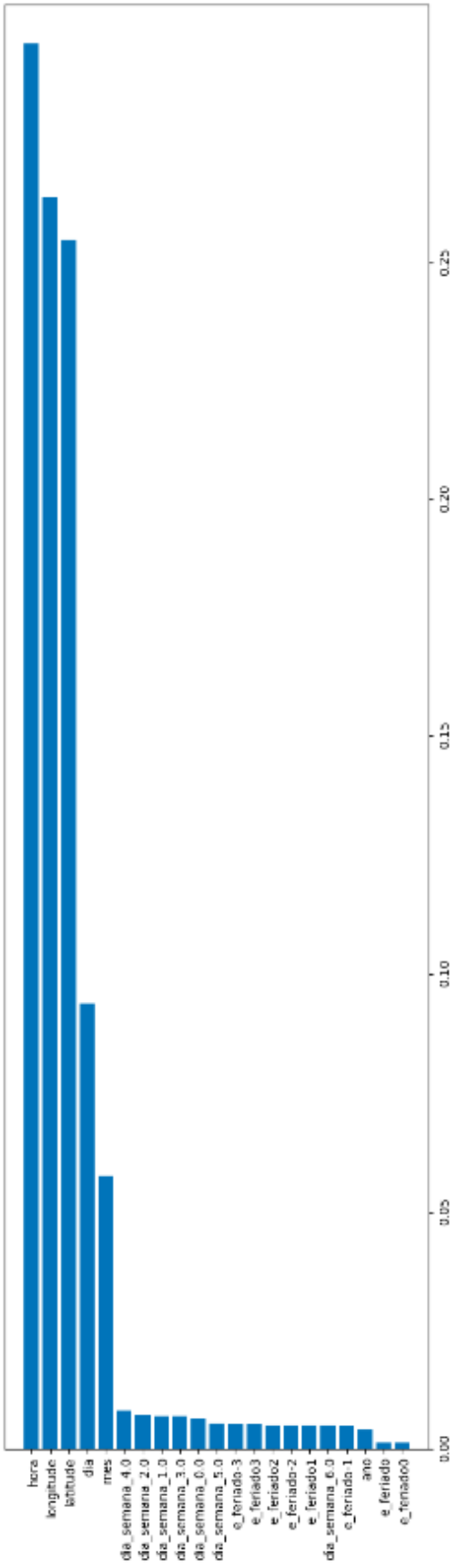


Figura A.2: feature importance.

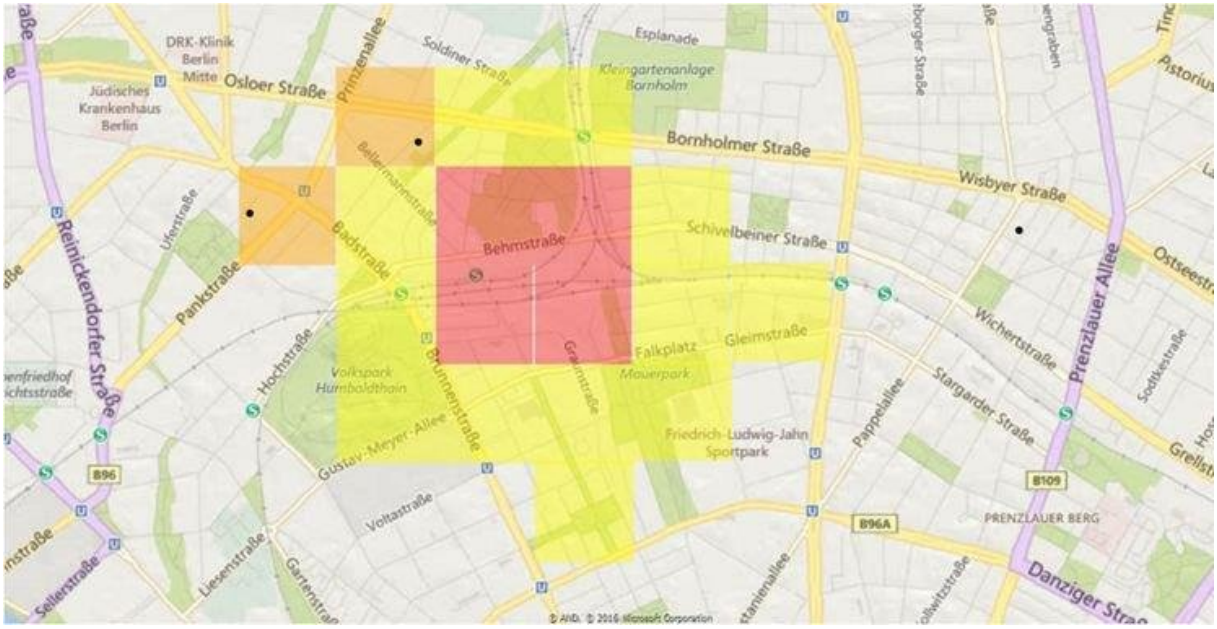


Figura A.3: Mapa com clusters preditivos no formato de Grid (Fonte: Próprio autor).

A.6 Aplicação de Redes Neurais Convolucionais na Formação de GRID

A formação das grades pode evoluir a partir da aplicação de técnicas de redes neurais convolucionais (CNNs) (LeCun et al., 1998, Goodfellow et al., 2016). De forma simplificada, a camada convolucional aplica um filtro (kernel) sobre a imagem de entrada, retornando uma imagem filtrada com valores diferentes dos pixels. Ou seja, ela utiliza duas imagens como entrada (a imagem real e o kernel) e retorna uma imagem como seu resultado. Este kernel atua deslizando sobre a imagem de entrada, aplicando sua função matemática e retornando a imagem modificada (Simonyan e Zisserman, 2014).

Com a aplicação de redes neurais convolucionais, é possível aprimorar a análise e a previsão de áreas de risco, permitindo identificar padrões mais complexos e sutis nos dados de crimes. As CNNs são especialmente eficazes em tarefas de detecção de padrões e reconhecimento de imagens, o que as torna ideais para analisar dados espaciais de crimes (He et al., 2016, Szegedy et al., 2015).

A utilização de CNNs no mapeamento preditivo de crimes pode oferecer várias vantagens, tais como:

- **Identificação de Padrões Complexos**: As CNNs podem detectar padrões que não seriam facilmente identificáveis por métodos tradicionais de análise de dados. Isso inclui a identificação de correlações espaciais e temporais que são críticas para prever áreas de risco (Goodfellow et al., 2016).

!ht]

Figura A.4: Mapa com clusters preditivos no formato de Grid com plotagem de setores (Fonte: Próprio autor).

- **Melhoria na Precisão das Previsões**: Ao analisar grandes volumes de dados históricos de crimes, as CNNs podem melhorar a precisão das previsões, destacando áreas que necessitam de maior atenção policial (He et al., 2016).

- **Visualização Aprimorada**: A capacidade das CNNs de processar e analisar dados visuais permite a criação de mapas de calor e outras representações visuais que facilitam a compreensão das áreas de risco (Goodfellow et al., 2016).

- **Automatização do Processo de Análise**: As CNNs podem automatizar grande parte do processo de análise, reduzindo a necessidade de intervenção manual e permitindo atualizações mais frequentes e precisas dos mapas de previsão (Simonyan e Zisserman, 2014, Szegedy et al., 2015).

O objetivo dessa abordagem é aprimora a precisão das previsões e também fornece ferramentas avançadas para a visualização e análise de dados, tornando o processo de policiamento mais eficiente e eficaz (Goodfellow et al., 2016, He et al., 2016).

A.6.1 Implementação de Redes Neurais Convolucionais

A aplicação de CNNs para melhorar a análise preditiva criminal envolve a construção e treinamento de um modelo convolucional que pode detectar padrões complexos nos dados. Abaixo, apresentamos um exemplo de código que ilustra como uma CNN pode ser aplicada a esse contexto.

A.6.2 Importação das Bibliotecas Necessárias

Primeiramente, importamos as bibliotecas necessárias para a construção e treinamento da rede neural convolucional, bem como para a manipulação e avaliação dos dados.

```
from sklearn.preprocessing import StandardScaler
import pandas as pd
import folium
from folium.plugins import HeatMap
import numpy as np
from sklearn.cluster import KMeans
import warnings
from sklearn.metrics import confusion_matrix
from sklearn.metrics import balanced_accuracy_score
from sklearn.metrics import accuracy_score
from sklearn import preprocessing
import matplotlib.pyplot as plt
from sklearn import preprocessing
from sklearn.model_selection import train_test_split
import seaborn as sns
import os
from sklearn.preprocessing import OneHotEncoder

from google.colab import drive
drive.mount('/content/drive')
```

Neste trecho, as bibliotecas ‘numpy‘ e ‘tensorflow‘ são utilizadas para operações matemáticas e construção da rede neural. ‘sklearn‘ fornece ferramentas para pré-processamento de dados e avaliação do modelo, enquanto ‘seaborn‘ e ‘matplotlib‘ são usadas para visualização.

A.6.3 Funções Auxiliares para Pré-Processamento de Dados

Além da normalização, a implementação de funções auxiliares são importantes no pré-processamento dos dados, permitindo que informações temporais e espaciais sejam extraídas e organizadas de forma eficiente. Essas funções são essenciais para o mapeamento de padrões e discretização dos dados.

A.6.3.1 Extração de Informações Temporais

As funções `get_hora` e `get_periodo` são utilizadas para extrair informações temporais de registros de horário. A função `get_hora` retorna a hora como

um número inteiro, enquanto `get_periodo` classifica os horários em períodos do dia, como manhã, tarde, noite ou madrugada. Essas informações podem ser incorporadas ao conjunto de dados para identificar padrões temporais de eventos criminais.

```
def get_hora(input: str) -> int:
    try:
        return int(input[0:2])
    except Exception:
        return -1

def get_periodo(input: str) -> str:
    try:
        hour = int(input[0:2])
        if hour <= 6:
            return 'Madrugada'
        elif hour <= 12:
            return 'Manha'
        elif hour <= 18:
            return 'Tarde'
        else:
            return 'Noite'
    except Exception:
        return 'Desconhecido'
```

Essas funções permitem o enriquecimento dos dados com categorias temporais, facilitando análises baseadas no horário dos eventos. Por exemplo, é possível identificar horários críticos para crimes em determinadas regiões.

A.6.3.2 Discretização de Valores Contínuos

A função `get_grid_pos` é utilizada para discretizar valores contínuos, como coordenadas geográficas, em uma grade regular. Essa discretização facilita a análise espacial, como a clusterização e a criação de mapas de calor.

```
def get_grid_pos(value: float, min_value: float, step: float) -> str:
    try:
        return np.round((value - min_value) / step)
    except Exception:
        return None
```

Ao transformar valores contínuos em posições discretas, é possível organizar os dados em uma estrutura uniforme, essencial para a aplicação de algoritmos como KMeans e para a criação de mapas de densidade.

A.6.4 Padronização dos Dados para Redes Neurais Convolucionais (CNNs)

A preparação dos dados para Redes Neurais Convolucionais (CNNs) envolve filtragem, enriquecimento e organização em um formato que permita a utiliza-

ção eficiente pelos modelos. As CNNs, originalmente projetadas para processar imagens, podem ser adaptadas para dados estruturados, representando-os em matrizes multi-dimensionais (ou "imagens") com diferentes canais de entrada.

A.6.4.1 Filtragem de Dados

Os dados foram inicialmente filtrados para incluir apenas as categorias de crimes mais relevantes, registrados na cidade de São Paulo, com coordenadas válidas (latitude e longitude) e dentro do período de 1ž de janeiro de 2022 a 1ž de março de 2023. Essa etapa reduz ruídos e foca a análise em dados significativos.

```
classes_crimes = ['FURTO - OUTROS', 'FURTO DE CARGA', 'FURTO DE VEÍCULO',
                  'ROUBO - OUTROS', 'ROUBO DE CARGA', 'ROUBO DE VEÍCULO']

df = df.loc[df['descricao'].isin(classes_crimes)]
df = df.loc[df['cidade'] == 'S.PAULO']
df = df.loc[df['latitude'] != 0.0]
df = df.loc[df['longitude'] != 0.0]
df = df.loc[(df['data'] >= '2022-01-01') & (df['data'] <= '2023-03-01')]
```

A.6.4.2 Enriquecimento dos Dados

A informação temporal foi enriquecida com atributos derivados, como o período do dia (madrugada, manhã, tarde ou noite), o dia, o mês e o dia da semana. Esses atributos são críticos para capturar padrões temporais.

```
df['data'] = pd.to_datetime(df['data'])
df['periodo'] = df['hora'].astype(str).apply(get_periodo)
df['dia'] = df['data'].dt.day
df['mes'] = df['data'].dt.month
df['dia_semana'] = df['data'].dt.weekday
```

A.6.4.3 Transformação em Matrizes Multi-Dimensionais

Para adaptar os dados a uma CNN, variáveis categóricas, como a descrição do crime, foram codificadas numericamente usando o *LabelEncoder*. Informações sobre o dia da semana foram transformadas em *one-hot encoding*, criando canais adicionais para essas categorias. Esses canais representam dimensões adicionais que as CNNs podem processar simultaneamente, como se fossem diferentes "camadas"de uma imagem.

```
le = preprocessing.LabelEncoder()
df['descricao'] = le.fit_transform(df['descricao'])
df = pd.get_dummies(df, columns=['dia_semana']).reset_index()
```

A.6.4.3.1 Canais e Dimensões dos Dados Os dados estruturados foram organizados em tensores multi-dimensionais, onde: - **Dimensão espacial**:

Representa as coordenadas geográficas (latitude e longitude). - **Canais adicionais**: Incluem informações temporais (período do dia, dia da semana) e categóricas (descrição do crime).

Assim permite que as CNNs tratem os dados como "imagens"multi-dimensionais, aproveitando sua capacidade de identificar padrões complexos em diferentes dimensões.

A.6.5 Construção dos Tensores para Redes Neurais Convolucionais

A etapa de construção dos tensores foi muito importante para transformar os dados pré-processados em um formato compatível com Redes Neurais Convolucionais (CNNs). Neste experimento, os dados estruturados foram organizados em uma matriz multidimensional (*tensor*) que captura as interações entre dimensões temporais, espaciais e categorias de crimes. Com essa organização foi possível que a CNN aprendesse padrões complexos e explorasse a relação entre essas dimensões.

A.6.5.1 Estrutura do Tensor

O tensor criado possui as seguintes dimensões:

$$\text{Tensor}[T, H, W, C]$$

Onde:

- T: Índice temporal, composto por combinações de datas e períodos do dia.
- H: Representa as posições de latitude (*latitude_grid_pos*).
- W: Representa as posições de longitude (*longitude_grid_pos*).
- C: Categorias de crimes, derivadas das descrições presentes no conjunto de dados.

Essa estrutura permite que o modelo utilize informações temporais e espaciais integradas para realizar predições mais precisas e robustas.

A.6.5.2 Preenchimento do Tensor

O preenchimento do tensor foi realizado iterando sobre os registros do conjunto de dados e associando as dimensões com seus valores correspondentes. O processo é exemplificado no código abaixo:

```
1 for idx, row in df_combined.iterrows():
2     row_date = row['data'].date()  # Extrair apenas a data
3     if row_date in distinct_data and row['periodo'] in distinct_periodo:
```

```
4           time_idx = distinct_data.tolist().index(row_date) * len(
                distinct_periodo) + \
5                       distinct_periodo.tolist().index(row['periodo'])
6           lat_idx = int(row['latitude_grid_pos'])
7           lon_idx = int(row['longitude_grid_pos'])
8           crime_idx = int(row['descricao'])
9
10          # Preencher o tensor
11          if 0 <= time_idx < tensor.shape[0] and \
12             0 <= lat_idx < tensor.shape[1] and \
13             0 <= lon_idx < tensor.shape[2] and \
14             0 <= crime_idx < tensor.shape[3]:
15              tensor[time_idx, lat_idx, lon_idx, crime_idx] += row['qtd']
16          else:
17              print(f"Índices fora do limite: time_idx={time_idx}, "
18                    f"lat_idx={lat_idx}, lon_idx={lon_idx}, crime_idx={
                          crime_idx}")
19      else:
20          print(f"Valor não encontrado: data={row['data']}, periodo={row['
                periodo']}")
```

A.6.5.3 Lógica do Preenchimento

1. Índice Temporal (T): Calculado combinando a data e o período do dia, permitindo capturar a sequência temporal dos eventos.

2. Índices Espaciais (H e W): Correspondem às posições de latitude e longitude discretizadas.

3. Índice de Categoria (C): Representa a descrição dos crimes, codificada em valores numéricos.

Os valores no tensor são incrementados com base no número de ocorrências (*qtd*) registradas para cada combinação de índices.

A.6.5.4 Formato Final do Tensor

- Dimensão Temporal (T): Número de combinações de datas e períodos do dia.

- Dimensões Espaciais (H e W): Grades discretizadas de latitude e longitude.

- Dimensão de Categorias (C): Número de categorias distintas de crimes.

O formato final é validado com:

```
1 print("Tensor shape:", tensor.shape)
```

A.6.6 Divisão de Dados: Conjuntos de Treino, Validação e Teste

A divisão dos dados é uma etapa essencial para garantir que o modelo seja treinado, ajustado e avaliado adequadamente. A estratégia adotada incluiu três conjuntos distintos:

- Treino (70%): Usado para ajustar os pesos do modelo.
- Validação (20%): Usado para monitorar o desempenho do modelo durante o treinamento e evitar *overfitting*.
- Teste (10%): Usado para avaliar o modelo em dados não vistos.

A.6.6.1 Normalização dos Dados

O tensor foi normalizado para garantir que os valores de entrada estivessem entre 0 e 1, o que melhora a estabilidade e a convergência do modelo:

```
tensor_normalized = tensor / tensor.max()
```

A.6.6.2 Divisão dos Conjuntos

A divisão foi realizada utilizando o método `train_test_split`, garantindo uma separação balanceada dos dados em conjuntos de treino, validação e teste.

```
from sklearn.model_selection import train_test_split

# Divisão em treino (70%), validação (20%) e teste (10%)
X_train, X_temp = train_test_split(tensor_normalized, test_size=0.3,
    random_state=42)
X_val, X_test = train_test_split(X_temp, test_size=0.33, random_state=42)

print(f"Treino: {X_train.shape}, Validação: {X_val.shape}, Teste: {X_test.
    shape}")
```

O resultado final foi:

- Treino: 70% do total de dados.
- Validação: 20%.
- Teste: 10%.

A.6.6.3 Dimensões para o Modelo

Com base no conjunto de treino, o formato de entrada para o modelo foi definido como:

```
input_shape = (X_train.shape[1], X_train.shape[2], X_train.shape[3])
```

Esse formato preserva as dimensões espaciais e as categorias de crimes, permitindo que a CNN aprenda padrões complexos ao longo do tempo e do espaço.

A.6.7 Aplicação na CNN

O tensor resultante foi utilizado como entrada para o modelo CNN, configurado com camadas convolucionais e de *pooling* para capturar padrões espaço-temporais. A seguir, o modelo foi treinado com os dados divididos, e os resultados confirmaram sua eficácia na análise preditiva de crimes.

A.6.8 Funções de Perda e sua Importância no Ajuste do Modelo

As funções de perda são componentes essenciais no treinamento de redes neurais, responsáveis por quantificar a diferença entre as predições do modelo (y_{pred}) e os valores reais (y_{true}). Elas guiam o ajuste dos pesos durante o aprendizado, garantindo que o modelo minimize erros e aprenda padrões relevantes nos dados. Neste trabalho, duas funções de perda foram exploradas: a *Weighted Loss* e a *Focal Loss*, ambas desempenhando papéis específicos no ajuste do modelo.

A.6.9 Weighted Loss

A *Weighted Loss* aplica penalizações diferenciadas aos erros com base no valor do dado verdadeiro (y_{true}). Essa abordagem é particularmente útil quando há regiões ou classes de maior importância nos dados, permitindo que o modelo foque nesses elementos durante o treinamento. A função é definida como:

$$\text{Loss} = w_0 \cdot y_{\text{true}} \cdot (y_{\text{pred}} - y_{\text{true}})^2 + w_1 \cdot (1 - y_{\text{true}}) \cdot (y_{\text{pred}} - y_{\text{true}})^2, \tag{A.1}$$

onde w_0 e w_1 são os pesos que controlam a penalização em regiões de maior e menor importância, respectivamente. Neste trabalho, foram definidos $w_0 = 3.0$ e $w_1 = 1.0$, enfatizando regiões de maior intensidade no y_{true}.

Na Figura A.5, observa-se que o modelo inicialmente gera uma saída uniforme, pois ainda não aprendeu a identificar padrões importantes. Isso é comum no início do treinamento, quando os pesos ainda não foram ajustados significativamente.

Na Figura A.6, nota-se uma evolução no aprendizado: o modelo começa a identificar padrões de maior intensidade, como pontos destacados. Esse progresso reflete o impacto das penalizações diferenciadas definidas pela *Weighted Loss*.

Por fim, na Figura A.7, o modelo demonstra a capacidade de prever padrões mais refinados e alinhados aos dados reais, indicando que a *Weighted Loss* cumpriu seu objetivo de guiar o aprendizado.

A.6.9.1 Focal Loss

A *Focal Loss* é projetada para lidar com cenários de desbalanceamento de classes, priorizando o aprendizado em amostras mais difíceis de classificar.

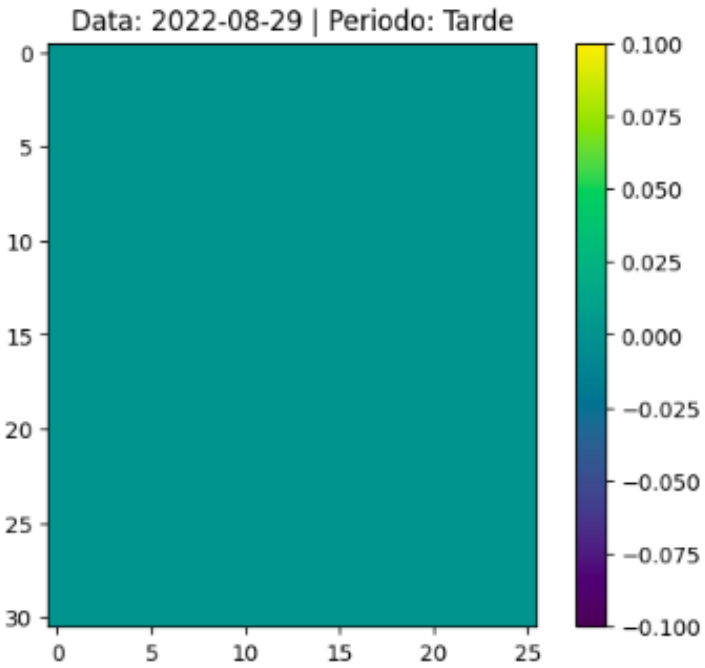


Figura A.5: Predição inicial do modelo. Fonte: Próprio autor

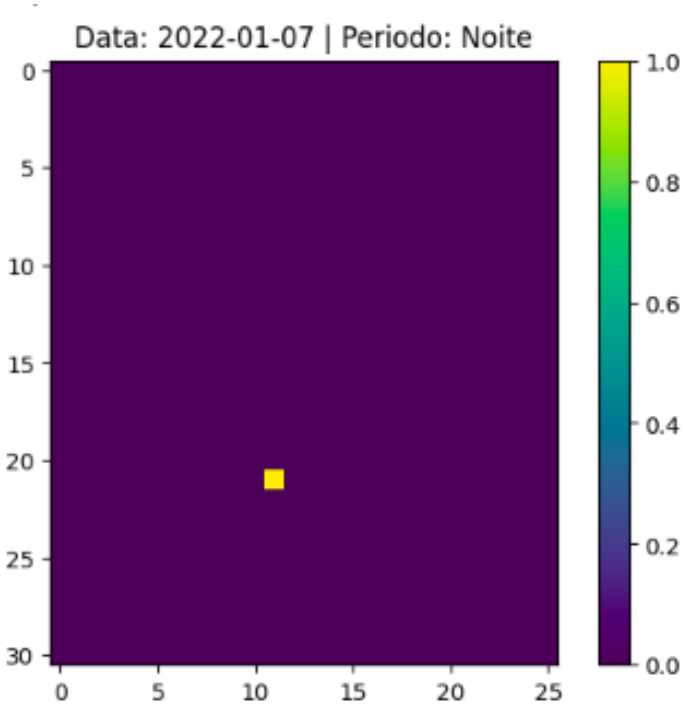


Figura A.6: Predição intermediária do modelo. Fonte: Próprio autor

Sua definição é:

$$\text{Loss} = -\alpha \cdot (1 - p_t)^{\gamma} \cdot \log(p_t), \tag{A.2}$$

onde:

- p_t é a probabilidade prevista para a classe correta.

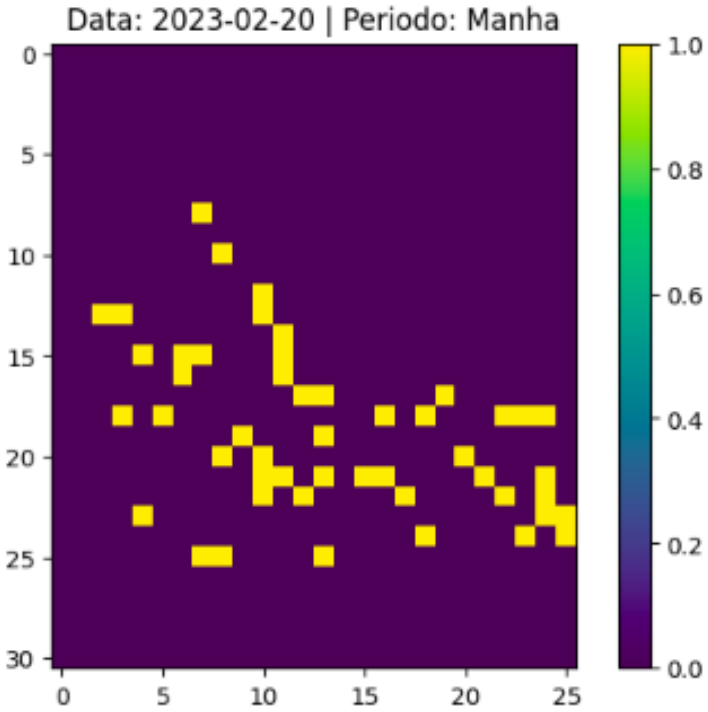


Figura A.7: Predição final do modelo. Fonte: Próprio autor

- α é um fator de ajuste que controla a importância relativa das classes.
- γ regula o foco nas amostras difíceis.

Embora a *Focal Loss* não tenha sido utilizada diretamente no treinamento final, sua análise destacou sua potencial aplicabilidade em problemas com desbalanceamento severo, pois ela reduz a penalização em amostras bem classificadas e concentra os esforços do modelo nas mais desafiadoras.

A.6.10 Impacto das Funções de Perda no Ajuste do Modelo

As funções de perda impactam diretamente o aprendizado do modelo ao ajustar os pesos das camadas. A *Weighted Loss* foi fundamental para orientar o modelo a capturar padrões relevantes, conforme ilustrado na Figura A.7. O uso de penalizações diferenciadas ajudou a priorizar áreas de maior relevância nos dados, enquanto reduziu o impacto de áreas menos significativas.

Essa abordagem garantiu que o modelo evoluísse de saídas uniformes (Figura A.5) para predições detalhadas (Figura A.7), destacando a importância de uma escolha cuidadosa da função de perda no desenvolvimento de modelos robustos.

A.6.11 Análise dos Dados de Treinamento e Resultados com poucas Épocas Inicialmente

Durante o treinamento do modelo, foram analisados os seguintes aspectos: o número de amostras utilizadas para treino e teste, o desempenho por época

em termos de *loss* e *MAE* (erro médio absoluto), e o impacto das funções de perda. Abaixo está um resumo dos dados observados:

- Conjunto de treino: 826 amostras de entrada com dimensões $(28, 31, 26, 1)$.
- Conjunto de teste: 44 amostras de entrada com dimensões $(28, 31, 26, 1)$.
- Saída esperada: Imagens com dimensões $(31, 26, 1)$.

O treinamento foi realizado por 5 épocas inicialmente, e as métricas de *loss* e *MAE* foram monitoradas tanto para os dados de treino quanto para validação. O gráfico da Figura A.8 ilustra a evolução dessas métricas ao longo das épocas.

```
(826, 28, 31, 26, 1)
(826, 31, 26, 1)
(44, 28, 31, 26, 1)
(44, 31, 26, 1)
Epoch 1/5
83/83 ━━━━━━━━━━━━━━━━━━━━ 533s 6s/step - loss: 0.1085 - mae: 0.1055 - val_loss: 0.2660 - val_mae: 0.0922
Epoch 2/5
83/83 ━━━━━━━━━━━━━━━━━━━━ 564s 6s/step - loss: 0.0713 - mae: 0.0290 - val_loss: 0.2660 - val_mae: 0.0922
Epoch 3/5
83/83 ━━━━━━━━━━━━━━━━━━━━ 573s 7s/step - loss: 0.0822 - mae: 0.0325 - val_loss: 0.2660 - val_mae: 0.0922
Epoch 4/5
83/83 ━━━━━━━━━━━━━━━━━━━━ 544s 7s/step - loss: 0.0738 - mae: 0.0298 - val_loss: 0.2660 - val_mae: 0.0922
Epoch 5/5
83/83 ━━━━━━━━━━━━━━━━━━━━ 542s 7s/step - loss: 0.0691 - mae: 0.0283 - val_loss: 0.2660 - val_mae: 0.0922
2/2 ━━━━━━━━━━━━━━━━━━━━ 5s 879ms/step - loss: 0.2643 - mae: 0.0916
Loss: 0.2659880518913269, MAE: 0.09217798709869385
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 255ms/step
```

Figura A.8: Evolução do *loss* e *MAE* durante as 5 primeiras épocas de treinamento do modelo. Fonte: Próprio autor

Análise dos Resultados

- Conjunto de treino: - A métrica de *loss* começou em 0.1085 na primeira época e apresentou uma redução progressiva, atingindo 0.0565 na última época. - O *MAE* reduziu de 0.1055 para 0.0242, indicando que o modelo conseguiu ajustar-se bem aos dados de treino.

- Conjunto de validação: - O *loss* e o *MAE* mantiveram-se constantes após a primeira época, em 0.2660 e 0.0922, respectivamente. Isso sugere que o modelo enfrentou dificuldades para generalizar os padrões aprendidos para os dados de validação.

Possíveis Interpretações

1. Estabilidade nas métricas de validação: - O *loss* e o *MAE* constantes no conjunto de validação indicam que o modelo pode estar superajustando os dados de treino, especialmente porque há uma discrepância na redução das métricas entre os conjuntos.

2. Evolução positiva no treino: - A redução do *loss* no treino demonstra que o modelo conseguiu aprender os padrões presentes no conjunto de treino, possivelmente devido à influência das funções de perda utilizadas.

3. Necessidade de regularização: - O modelo poderia se beneficiar de técnicas como *dropout*, *early stopping* ou maior diversificação dos dados de treino para melhorar a capacidade de generalização.

4. Tamanho dos conjuntos: - O número de amostras no conjunto de teste é relativamente pequeno (44 amostras), o que pode dificultar a avaliação da generalização do modelo.

Conclusão e Próximos Passos

Os resultados indicam que o modelo conseguiu aprender os padrões do conjunto de treino, mas apresenta limitações na validação. Como próximos passos, sugere-se:

- Aumentar o conjunto de validação para reduzir o viés na avaliação.
- Implementar regularização para evitar sobreajuste.
- Avaliar o impacto de diferentes funções de perda no desempenho do modelo.

A.6.12 Análise Binária das Previsões e Avaliação do Modelo

Nesta subseção, apresentamos o processo de análise das previsões realizadas pelo modelo, considerando uma abordagem binária para classificar as regiões geográficas com alta ou baixa probabilidade de ocorrência de crimes. As previsões do modelo foram ajustadas para destacar áreas de maior risco ($\geq 90\%$) e descartar regiões de baixa probabilidade ($< 70\%$). Em seguida, comparamos as imagens previstas com os dados reais e avaliamos os erros absolutos.

A.6.12.1 Normalização e Ajuste Binário da Imagem Prevista

Após o modelo realizar a predição, a imagem prevista (`predicted_image`) contém valores contínuos entre 0 e 1, representando a probabilidade de ocorrência de crimes. Para aplicar a análise binária, os valores foram ajustados conforme os seguintes critérios:

- Células com probabilidade maior ou igual a 90% foram classificadas como 1 (alta probabilidade de crime).
- Células com probabilidade menor que 70% foram classificadas como 0 (baixa probabilidade de crime).

O código abaixo ilustra o processo de ajuste binário:

```
# Criar uma cópia da imagem prevista
predicted_image_norm = predicted_image.copy()

# Ajustar os valores com base nos limiares binários
predicted_image_norm[predicted_image_norm >= 0.9] = 1
predicted_image_norm[predicted_image_norm < 0.7] = 0
```

Esse ajuste destaca apenas as áreas de maior interesse, simplificando a análise e a interpretação dos resultados.

A.6.12.2 Cálculo do Erro Absoluto na Análise Binária

Para avaliar o desempenho do modelo, comparamos a imagem prevista ajustada (`predicted_image_norm`) com a imagem real (`real_image`). O erro absoluto foi calculado pixel a pixel, refletindo as discrepâncias entre previsão e realidade.

```
# Selecionar a imagem real
real_image = y_test[0].squeeze()

# Calcular o erro absoluto
error_image = np.abs(predicted_image_norm - real_image)
```

Os valores na matriz `error_image` indicam:

- 0: Previsões corretas (concordância entre previsão e realidade).
- 1: Diferenças significativas (erro na previsão).

A.6.12.3 Visualização dos Resultados da Análise Binária

A seguir, apresentamos as imagens da análise binária, incluindo a imagem prevista ajustada, a imagem real e o erro absoluto:

```
# Plotar as imagens lado a lado
plt.figure(figsize=(12, 4))

# Imagem prevista ajustada
plt.subplot(1, 3, 1)
plt.imshow(predicted_image_norm, cmap='viridis', origin='upper')
plt.title("Imagem Prevista (Binária)")
plt.axis("off")

# Imagem real
plt.subplot(1, 3, 2)
plt.imshow(real_image, cmap='viridis', origin='upper')
plt.title("Imagem Real")
plt.axis("off")

# Erro absoluto
plt.subplot(1, 3, 3)
```

```
plt.imshow(error_image, cmap='viridis', origin='upper')
plt.title("Erro Absoluto")
plt.axis("off")

plt.show()
```

A.6.12.4 Resultados Ilustrativos

A Figura A.9 apresenta as imagens da análise binária. A imagem prevista binária destaca as regiões com alta probabilidade de crimes, enquanto a imagem real serve como referência para avaliação. A imagem do erro absoluto mostra as discrepâncias pixel a pixel entre previsão e realidade.

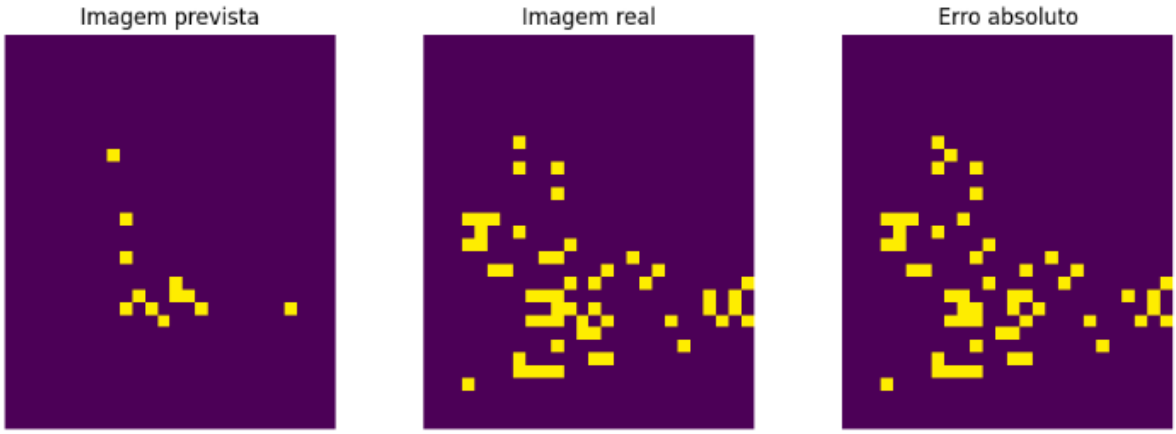


Figura A.9: Comparação entre a imagem prevista (binária), a imagem real e o erro absoluto. Fonte: Próprio autor

A.6.12.5 Interpretação dos Resultados

Os resultados apresentados na Figura A.9 mostram um padrão interessante na previsão da rede neural. A análise dos resultados da CNN revelou que, em um primeiro momento, houve um número reduzido de pontos de previsão com alta probabilidade ($\geq 90\%$). Esse comportamento pode ser explicado por dois fatores principais:

1. **Foco nos centróides das regiões de risco**: A rede neural pode estar identificando apenas os pontos centrais das áreas com maior incidência criminal, sem considerar a influência da vizinhança. Como consequência, as previsões se concentram em poucos locais e não refletem toda a dispersão observada nos dados reais.

2. **Limitação no tempo de treinamento**: Devido às restrições computacionais do ambiente de experimentação (Google Colab), o modelo foi treinado em apenas **5 épocas** e com um período de análise limitado a **28 dias subdivididos em quatro turnos** (manhã, tarde, noite e madrugada). Em um ambiente com maior capacidade computacional, seria possível aumentar o número de

épocas e expandir a base temporal, o que poderia melhorar a generalização do modelo e refinar as previsões.

Para aprimorar a eficácia da predição criminal, sugere-se a combinação do modelo convolucional com outros classificadores que levem em conta a vizinhança dos eventos. Um candidato promissor é o **K-Nearest Neighbors (KNN)**, que pode contribuir para a dispersão mais realista das previsões ao considerar a densidade dos pontos próximos.

A aplicação conjunta dessas técnicas pode ser estruturada da seguinte forma:

- A **rede neural** pode ser utilizada como um modelo primário de detecção, identificando os **pontos centrais das áreas críticas**.
- O **KNN** pode ser aplicado posteriormente para **expandir a previsão**, distribuindo os pontos de risco de forma mais alinhada à distribuição observada dos crimes.

Essa abordagem híbrida permitiria capturar tanto os padrões estruturais dos crimes, detectados pela CNN, quanto a distribuição espacial detalhada, ajustada pelo KNN.

A.7 Análise dos resultados aferidos com a implementação da CNN na Predição Criminal

Este experimento investigou o uso de Redes Neurais Convolucionais 3D (CNNs 3D) na previsão de ocorrências criminais, com foco na identificação de padrões espaço-temporais em uma grade geográfica representando regiões monitoradas. A metodologia explorou a capacidade de CNNs 3D de processar grandes volumes de dados espaço-temporais, respeitando as limitações computacionais do ambiente Google Colab, e a escolha de uma janela temporal de 28 dias, equivalente a 4 semanas, foi justificada pela possibilidade de capturar padrões temporais básicos e sazonalidades semanais relevantes para a análise criminal.

A.7.1 Metodologia e Estrutura do Experimento

A.7.1.1 Organização dos Dados

Os dados consistem em imagens bidimensionais geradas a partir de registros criminais em uma grade geoespacial definida por latitude e longitude. Cada pixel de uma imagem representa o número de ocorrências em uma célula geográfica específica em um determinado período do dia. A granularidade temporal dos dados foi dividida em quatro períodos diários: madrugada, manhã, tarde e noite, totalizando 28 imagens para um intervalo de 7 dias.

A organização final dos dados incluiu a formação de sequências temporais para entrada no modelo:

- Cada sequência de entrada (X) foi composta por 28 imagens consecutivas;
- A saída correspondente (y) foi definida como a imagem representando o próximo período após a sequência.

O processo de preparação dos dados foi automatizado utilizando o seguinte código:

```
X, y = [], []
for i in range(len(imagens) - 28):
    X.append(imagens[i:i + 28])  # Sequências de 28 imagens
    y.append(imagens[i + 28])    # Próxima imagem
```

Essa estrutura permitiu capturar dependências temporais de curto prazo (semanais) e explorar como os padrões criminais evoluem ao longo dos períodos do dia.

A.7.1.2 Arquitetura da Rede Neural

A arquitetura da CNN 3D foi projetada para processar sequências espaço-temporais e modelar as interdependências espaciais e temporais. A estrutura incluiu:

- `Conv3D`: Camadas convolucionais tridimensionais para extração de características espaço-temporais;
- `MaxPooling3D`: Camadas de redução de dimensionalidade, diminuindo a carga computacional e o risco de overfitting;
- `Dense`: Uma camada densa para aprendizado de alto nível;
- `Reshape`: Reconstrução do formato original da imagem prevista.

O modelo foi implementado como segue:

```
model = models.Sequential([
    layers.Conv3D(64, (3, 3, 3), activation='relu', padding="same",
                  input_shape=(28, 31, 26, 1)),
    layers.MaxPooling3D((2, 2, 1)),
    layers.Conv3D(128, (3, 3, 3), activation='relu', padding="same"),
    layers.Flatten(),
    layers.Dense(1024, activation='relu'),
    layers.Dense(np.prod((31, 26, 1)), activation='sigmoid'),
    layers.Reshape((31, 26, 1))
])
```

A.7.1.3 Treinamento e Avaliação

O modelo foi treinado por 5 épocas, com divisão de dados em 80% para treinamento e 20% para teste, utilizando uma função de perda ponderada para lidar com possíveis desequilíbrios nos dados. A métrica de avaliação incluiu o erro absoluto pixel a pixel entre a imagem prevista e a real:

```
def weighted_loss(y_true, y_pred):
    weight_0 = 3.0
    weight_1 = 1.0
    loss = weight_0 * y_true * tf.square(y_pred - y_true) + \
           weight_1 * (1 - y_true) * tf.square(y_pred - y_true)
    return tf.reduce_mean(loss)
```

A.7.2 Comparação utilizando-se dos Classificadores Tradicionais

Para avaliar a eficácia da CNN 3D, comparamos os resultados com modelos clássicos de aprendizado de máquina. Os classificadores analisados foram:

- `DecisionTreeClassifier`;
- `LogisticRegression`;
- `KNeighborsClassifier`;
- `RandomForestClassifier`;
- `GradientBoostingClassifier`;
- `GaussianNB`.

A.7.2.1 Logistic Regression

O modelo apresentou o melhor desempenho entre todos os classificadores, com **Accuracy** e **Balanced Accuracy** de 97.89%. Os valores de **Precision**, **Recall** e **F1-Score** foram altos e equilibrados para ambas as classes (0.0 e 1.0), destacando-se como um modelo eficaz para o problema. Esse classificador é indicado para problemas onde a separação entre as classes pode ser explicada por uma relação linear.

A.7.2.2 K-Nearest Neighbors (KNN)

Embora tenha apresentado um bom desempenho geral, com **Accuracy** de 95.47%, ficou abaixo da Logistic Regression e do Gradient Boosting. O modelo demonstrou um **Recall** ligeiramente inferior para a classe 0.0 (93%), indicando dificuldades em identificar corretamente todas as amostras dessa classe. KNN é sensível à escolha do parâmetro k e pode ser potencialmente afetado por outliers.

A.7.2.3 Random Forest

O modelo apresentou resultados consistentes, com **Accuracy** de 96.37% e **Balanced Accuracy** equivalente. Demonstrou equilíbrio entre **Precision** e **Recall** para ambas as classes (0.0 e 1.0), evidenciando sua robustez em dados com maior complexidade. Apesar de ser resiliente a overfitting, foi ligeiramente superado pelo Gradient Boosting.

A.7.2.4 Gradient Boosting

Foi o segundo melhor modelo, com **Accuracy** de 96.98% e **Balanced Accuracy** equivalente. Apresentou resultados uniformes em todas as métricas, com **Precision**, **Recall**, e **F1-Score** de 97% para ambas as classes. Esse modelo é indicado para problemas complexos, onde há maior dependência de relações não lineares entre as features.

A.7.2.5 Gaussian Naive Bayes

O modelo apresentou um desempenho insatisfatório, com **Accuracy** de apenas 54.38% e **Balanced Accuracy** de 54.47%. A classe 0.0 teve um **Recall** muito baixo (25%), indicando dificuldade em prever corretamente essa classe. Este classificador é inadequado para os dados avaliados, possivelmente devido à suposição de independência entre as features, que não se aplica ao problema.

A.7.2.6 Conclusão Geral

A **Logistic Regression** apresentou o melhor desempenho geral, sendo recomendada para este conjunto de dados. **Gradient Boosting** e **Random Forest** são alternativas viáveis, apresentando resultados robustos e consistentes. Modelos como **KNN** são aceitáveis, mas apresentam limitações em relação a dados complexos e dependem de ajustes no parâmetro k. O modelo **Gaussian Naive Bayes** não é recomendado, devido ao baixo desempenho em problemas deste tipo.

A.7.2.7 Imagem Ilustrativa dos Resultados

Para complementar a análise, apresentamos abaixo uma imagem que ilustra o comportamento do erro absoluto entre a imagem prevista e a imagem real pela CNN 3D. Essa visualização ajuda a entender a distribuição do erro no modelo:

A.7.3 Visualização em Matriz de Confusão

Como exemplo, a matriz de confusão é visualizada para facilitar a interpretação dos resultados.

```
# Plotar a matriz de confusão
plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues')
plt.title('Matriz de Confusão')
```

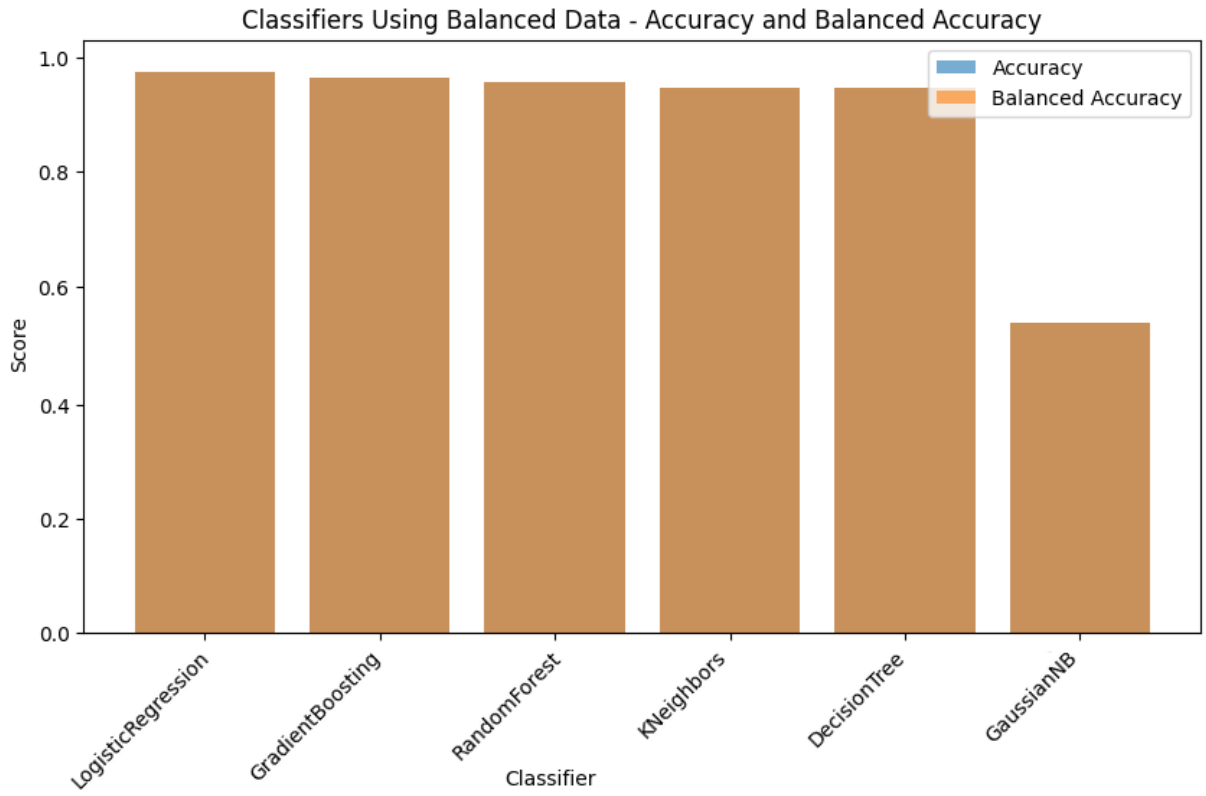


Figura A.10: Comparação entre a imagem prevista, a imagem real e o erro absoluto. Fonte: Próprio autor

```
plt.xlabel('Previsto')
plt.ylabel('Real')
plt.show()
```

'seaborn.heatmap' é usado para criar um gráfico da matriz de confusão, que ajuda a visualizar a performance do modelo na classificação das diferentes classes de crimes.

Conclui-se que abordagem de usar CNNs pode melhorar muito a precisão e a eficácia das previsões criminais, permitindo uma análise mais sofisticada e detalhada dos dados.

A.7.3.1 Análise da Matriz de Confusão

A matriz abaixo detalha o comportamento do modelo ao classificar as amostras no conjunto de teste.

A.7.3.2 Valores na Matriz

- **Classe 0.0 (Verdadeiro Negativo e Falso Positivo)**:
  - **159**: Amostras da classe 0.0 corretamente classificadas como 0.0 (**Verdadeiro Negativo**).
  - **7**: Amostras da classe 0.0 incorretamente classificadas como 1.0 (**Falso Positivo**).

- **Classe 1.0 (Falso Negativo e Verdadeiro Positivo)**:
  - **165**: Amostras da classe 1.0 corretamente classificadas como 1.0 (**Verdadeiro Positivo**).
  - **0**: Amostras da classe 1.0 incorretamente classificadas como 0.0 (**Falso Negativo**).

A.7.3.3 Desempenho Geral

O classificador **Logistic Regression** apresentou um excelente desempenho geral:

- **Precisão (Precision)**:
  - Classe 0.0: 1.00 (100%).
  - Classe 1.0: ≈ 0.96 (96%).
- **Revocação (Recall)**:
  - Classe 0.0: ≈ 0.96 (96%).
  - Classe 1.0: 1.00 (100%).
- **Acurácia Geral (Accuracy)**: 97.89%.

A.7.3.4 Conclusão

O modelo foi eficiente para prever ambas as classes, com poucos erros. Ele classificou corretamente **97.89%** das amostras no conjunto de teste. Para a classe 1.0, o modelo obteve um desempenho perfeito, sem falsos negativos. No entanto, foram observados **7 falsos positivos** na classe 0.0. Esses resultados tornam o **Logistic Regression** uma excelente escolha para o problema abordado.

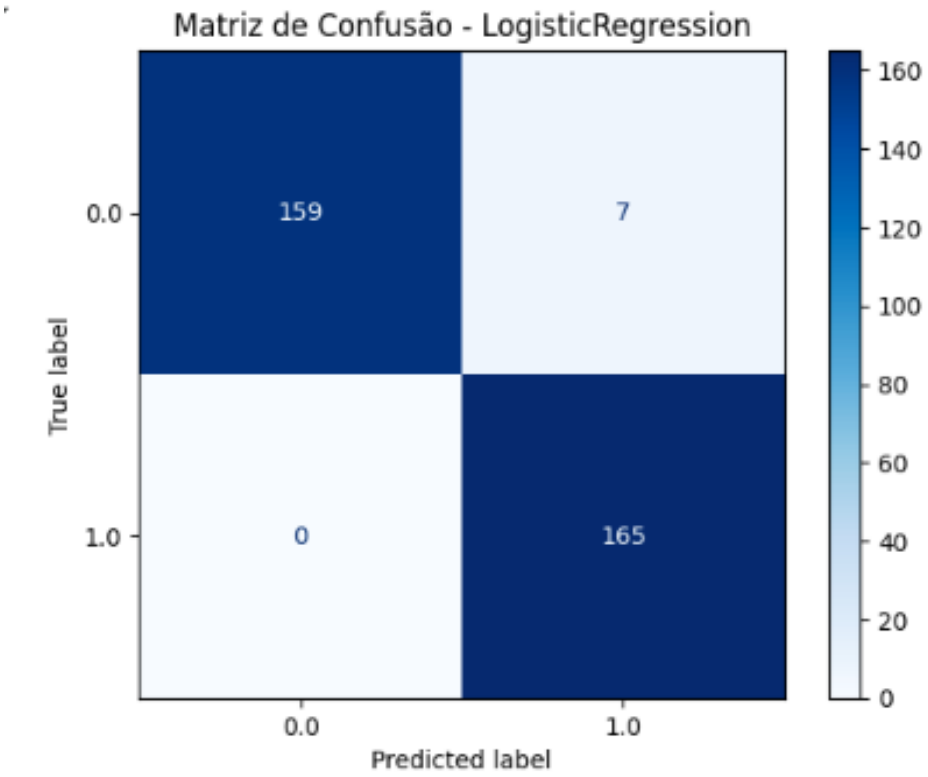


Figura A.11: Matriz de Confusão do Classificador Logistic Regression. Fonte: Próprio autor

Por causa da arquitetura da rede neural convolucional, é possível reduzir o número de pesos necessários para processar uma imagem como entrada (Krizhevsky et al., 2009). Utilizando redes neurais convolucionais (CNNs), o número de pesos é significativamente reduzido.

Isso ocorre porque, através do filtro (kernel), um neurônio de uma camada posterior está conectado a uma pequena região da imagem na camada anterior. Essa abordagem não só diminui o número total de pesos, mas também conserva as relações espaciais na imagem, preservando as posições relativas dos pixels. A capacidade das CNNs de preservar essas relações espaciais enquanto reduz o número de pesos é crucial para a análise eficiente de imagens e a detecção de padrões complexos, como os necessários para prever áreas de risco em mapas preditivos de crimes (LeCun et al., 1998).

Além de melhorar a eficiência computacional, as CNNs permitem uma análise mais detalhada e precisa dos dados espaciais, resultando em previsões mais acuradas e úteis para as operações de policiamento preditivo. Assim, a integração dessas técnicas avançadas pode potencialmente transformar o campo da previsão de crimes, oferecendo novas oportunidades para a prevenção e a segurança pública (Goodfellow et al., 2016).

A.7.4 Limitações

Apesar de promissor, o experimento teve limitações decorrentes do ambiente computacional:

- O uso do Google Colab restringiu a série histórica a 28 dias devido às limitações de memória e processamento;
- Não foram incorporadas variáveis contextuais adicionais, como fatores socioeconômicos ou ambientais, que poderiam enriquecer a modelagem.

A.8 Apresentação de técnicas Geoespaciais de K-means e Risk Terrain Modeling que podem ser integradas ao resultado da Predição Criminal

A previsão criminal é uma área crescente de estudo que visa usar dados históricos e técnicas analíticas para prever onde futuros crimes provavelmente ocorrerão. Duas abordagens populares nesta área são o K-means clustering e o Risk Terrain Modeling (RTM). Nesta seção, exploramos como essas técnicas podem ser aplicadas à predição criminal, utilizando dados geoespaciais para identificar padrões e áreas de risco.

A.8.1 K-means Clustering

O K-means clustering é um método de aprendizado não supervisionado usado para dividir um conjunto de dados em K clusters distintos. Cada ponto de dado é atribuído ao cluster cujo centroide é o mais próximo. Este método pode ser aplicado à predição criminal agrupando incidentes criminais em áreas específicas, permitindo a identificação de zonas de alta criminalidade.

A.8.1.1 Algoritmo K-means

O algoritmo K-means segue os seguintes passos:

1. Selecionar K centróides iniciais aleatoriamente.
2. Atribuir cada ponto de dado ao centróide mais próximo, formando K clusters.
3. Recalcular os centróides com base nos pontos atribuídos a cada cluster.
4. Repetir os passos 2 e 3 até que os centróides não mudem significativamente ou um número máximo de iterações seja alcançado.

A.8.1.2 Aplicação ao Dataset Criminal

Para aplicar K-means ao nosso dataset criminal, seguimos os passos do algoritmo, utilizando as coordenadas geoespaciais dos incidentes criminais como entradas. A escolha do valor de K pode ser feita através de métodos como o "cotovelo"ou a "silhueta".

```
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
import geopandas as gpd

# Carregar os dados
df = pd.read_csv('crime_data.csv')

# Filtrar dados geoespaciais
df = df[['longitude', 'latitude']]

# Remover valores nulos
df = df.dropna()

# Configurar o K-means
kmeans = KMeans(n_clusters=5, random_state=0)

# Ajustar o modelo
kmeans.fit(df)

# Adicionar os rótulos dos clusters ao dataframe
df['cluster'] = kmeans.labels_

# Plotar os clusters
plt.scatter(df['longitude'], df['latitude'], c=df['cluster'])
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.title('Clusters de Criminalidade')
plt.show()
```

A.8.1.3 Resultado

A figura abaixo mostra os clusters de criminalidade identificados pelo K-means, com os centroides dos clusters destacados em vermelho.

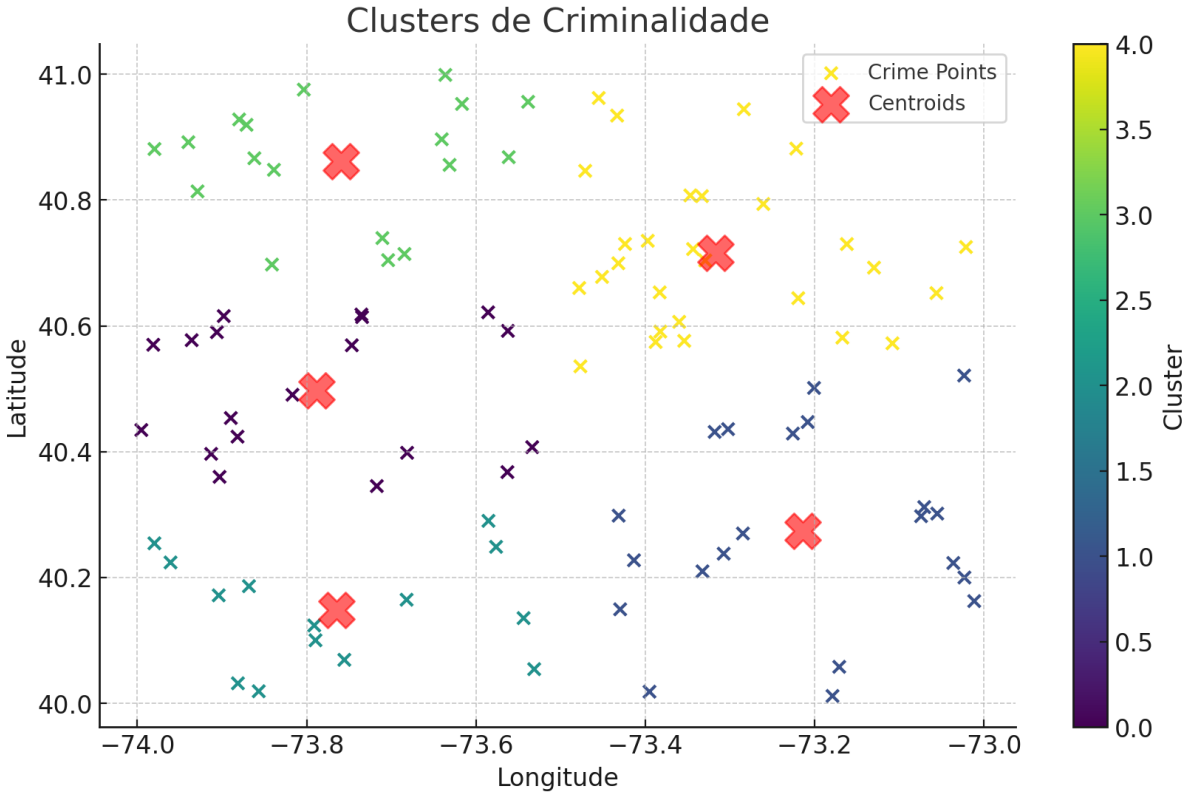


Figura A.12: Clusters de Criminalidade. Fonte: Próprio Autor

A.8.1.4 Interpretação da Imagem

A imagem acima mostra os clusters de criminalidade em uma área geográfica. Cada ponto colorido representa uma ocorrência de crime, agrupada em um dos cinco clusters distintos determinados pelo algoritmo K-means. As cores diferentes indicam a qual cluster cada ponto de crime pertence.

A.8.1.5 Legenda

- **Pontos coloridos:** Representam os incidentes de crimes, agrupados em diferentes clusters.
- **Centroides (marcados com 'X' vermelhos):** Representam o centro de cada cluster.

A.8.1.6 Interpretação

- **Pontos de Crimes:** Os pontos coloridos na imagem mostram a distribuição geográfica dos crimes. Crimes que estão próximos uns dos outros são agrupados no mesmo cluster.
- **Centroides:** Os centroides são os pontos centrais dos clusters, indicados pelos 'X' vermelhos. Estes são os locais sugeridos para posicionar as viaturas de polícia para otimizar a cobertura de áreas com alta incidência de crimes.

A.8.1.7 Posicionamento das Viaturas

As viaturas devem ser posicionadas nos locais marcados pelos centroides (X vermelhos) para cobrir melhor as áreas com maior concentração de crimes. Ao fazer isso, a polícia pode responder mais rapidamente a incidentes e aumentar a presença em locais críticos, potencialmente prevenindo novos crimes. Para maximizar a cobertura e a resposta rápida às áreas com maior ocorrência de crimes, as viaturas policiais devem ser posicionadas próximas aos centroides dos clusters (pontos vermelhos). Isso garante que a presença policial está centralizada em áreas com alta densidade criminal, facilitando a prevenção e a resposta rápida a incidentes.

A.8.2 Risk Terrain Modeling (RTM)

O Risk Terrain Modeling é uma técnica que identifica fatores ambientais associados ao risco de crime. Diferente do K-means, que agrupa incidentes criminais, o RTM analisa características espaciais que podem contribuir para a ocorrência de crimes, como a proximidade a bares, escolas ou áreas mal iluminadas.

A.8.2.1 Etapas do RTM

As etapas para realizar um RTM incluem:

1. Identificar fatores de risco relevantes (e.g., presença de bares, escolas).
2. Mapear esses fatores em um sistema de informações geográficas (SIG).
3. Combinar os mapas de fatores de risco para criar um mapa de risco agregado.
4. Validar o mapa de risco com dados de crime observados.

A.8.2.2 Aplicação ao Dataset Criminal

Ao aplicar o RTM ao nosso dataset criminal, mapeamos vários fatores de risco e combinamos esses mapas para criar um modelo de risco espacial. Este modelo é então usado para prever áreas de alta criminalidade, permitindo intervenções proativas por parte das forças de segurança.

```
import geopandas as gpd
from shapely.geometry import Point

# Carregar os dados de crime e fatores de risco
crime_data = pd.read_csv('crime_data.csv')
risk_factors = pd.read_csv('risk_factors.csv')

# Criar GeoDataFrames
```

```
crime_gdf = gpd.GeoDataFrame(crime_data, geometry=gpd.points_from_xy(
    crime_data.longitude, crime_data.latitude))
risk_gdf = gpd.GeoDataFrame(risk_factors, geometry=gpd.points_from_xy(
    risk_factors.longitude, risk_factors.latitude))

# Definir o sistema de coordenadas
crime_gdf = crime_gdf.set_crs('EPSG:4326')
risk_gdf = risk_gdf.set_crs('EPSG:4326')

# Criar um mapa de calor dos crimes
crime_heatmap = crime_gdf.geometry.apply(lambda x: x.buffer(0.01)).
    unary_union

# Criar mapas de calor para cada fator de risco
risk_heatmaps = {factor: risk_gdf[risk_gdf['factor'] == factor].geometry.
    apply(lambda x: x.buffer(0.01)).unary_union
                 for factor in risk_gdf['factor'].unique()}

# Combinar os mapas de calor dos fatores de risco
combined_risk = gpd.GeoSeries(risk_heatmaps.values()).unary_union

# Plotar os mapas
fig, ax = plt.subplots(figsize=(10, 10))
gpd.GeoSeries(crime_heatmap).plot(ax=ax, color='red', alpha=0.5)
gpd.GeoSeries(combined_risk).plot(ax=ax, color='blue', alpha=0.5)
plt.title('Mapa de Risco Combinado')
plt.show()
```

Podemos observar que K-means clustering e o Risk Terrain Modeling podem oferecer alguns benefícios para complementar a predição criminal. Enquanto o K-means é eficaz em identificar padrões e clusters de crimes, o RTM fornece uma compreensão mais profunda dos fatores ambientais que contribuem para o risco de crime. Combinando essas técnicas, é possível implementar estratégias de localização de máxima cobertura conforme a predição criminal acontece, buscando evitar o surgimento das ocorrências.

A.9 Pós-processamento

O pós-processamento é uma etapa crucial após a aplicação de algoritmos de aprendizado de máquina, especialmente devido à presença de ruídos e inconsistências nos dados processados. Antes que a extração de conhecimento possa ser efetivamente realizada, diversos procedimentos de pré-processamento são necessários para limpar, transformar e organizar os dados, garantindo que os algoritmos de aprendizado possam operar de forma eficaz.

Após a fase de mineração, os modelos resultantes, como árvores de decisão, conjuntos de regras ou topologias de redes neurais, podem não ser

adequados para a visualização direta dos resultados, devido à complexidade ou à presença de padrões irrelevantes. Para melhorar a compreensão desses modelos, é essencial realizar um pós-processamento (Bruha e Famili, 2000). Essa fase inclui procedimentos como filtragem de regras, integração e visualização de conhecimento, todos com o objetivo de refinar o conhecimento obtido e corrigir erros ou informações imprecisas.

O pós-processamento pode ser dividido nas seguintes categorias:

- **Filtro de Conhecimento**: Algoritmos indutivos, como árvores de decisão e regras, frequentemente geram folhas correspondentes a poucos exemplos do conjunto de treinamento. Esse fenômeno ocorre porque os algoritmos buscam ser o mais consistente possível com o conjunto de treinamento. Métodos para evitar isso incluem a poda de árvores e a truncagem de regras de decisão, que removem essas folhas problemáticas.

- **Interpretação e Explicação**: Após a obtenção do conhecimento pelos algoritmos, é crucial utilizá-lo de forma eficaz. Algumas técnicas podem não apresentar resultados de forma clara para o usuário final. Documentação, visualização e integração com sistemas existentes são abordagens que podem facilitar a legibilidade do conhecimento gerado.

- **Avaliação**: Esta fase é executada após a geração das hipóteses pelos algoritmos. Diversas métricas podem ser usadas para avaliar o desempenho do processo, incluindo precisão na classificação, legibilidade e complexidade computacional.

- **Integração de Conhecimento**: Sistemas tradicionais de apoio à decisão geralmente utilizam um único algoritmo de aprendizado. No entanto, métodos mais recentes combinam vários modelos de aprendizado. A integração desses modelos deve ser feita de maneira a evitar conflitos e incluir métodos de visualização e interpretação para apresentar os resultados produzidos.

Por meio desses métodos, é possível melhorar a precisão e a interpretação dos resultados, tornando os sistemas mais amigáveis e compreensíveis para o usuário final.

A.10 Considerações Finais

Neste capítulo, abordamos a predição criminal através da aplicação de técnicas avançadas de machine learning e análise espacial. Iniciamos o processo com a importação e preparação dos dados, incluindo a limpeza e a criação de variáveis adicionais, como a inclusão de feriados e variáveis de atraso, o que enriqueceu o conjunto de dados para melhor capturar as nuances dos

padrões de criminalidade. Em seguida, selecionamos e configuramos um conjunto diversificado de algoritmos de classificação, incluindo `Decision Tree`, `Logistic Regression`, `K-Nearest Neighbors`, `Random Forest`, `Gradient Boosting` e `Naive Bayes`, para prever o tipo de crime com base em várias características espaciais e temporais.

Os modelos foram treinados e avaliados com métricas como a matriz de confusão, a acurácia balanceada e a acurácia geral. A análise dos resultados revelou que o `Random Forest` e o `Gradient Boosting` mostraram os melhores desempenhos gerais, com acurácias de 64.2% e 64.1%, respectivamente. Esses modelos apresentaram um equilíbrio razoável na classificação das diferentes classes, embora ainda houvesse desafios na identificação precisa de categorias menos representadas. Em contraste, o `GaussianNB` apresentou a menor acurácia geral, indicando limitações na suposição de independência entre variáveis para este contexto específico.

Além disso, a visualização da matriz de confusão para cada classificador permitiu uma compreensão mais profunda das áreas de dificuldade dos modelos, enquanto a análise da importância das features destacou quais variáveis desempenham papéis cruciais nas previsões. A combinação dessas análises e técnicas sugere que a abordagem utilizada pode ser eficaz na identificação de hotspots de criminalidade. Essa eficácia pode contribuir significativamente para uma melhor alocação de recursos de segurança pública, promovendo intervenções mais direcionadas e potencialmente mais eficazes.

Adicionalmente, aplicamos técnicas de K-means e Risk Terrain Modeling (RTM) para identificar clusters de criminalidade e locais estratégicos para a colocação de viaturas de polícia. Através do K-means, determinamos os centroides dos clusters de criminalidade, que representam os pontos centrais de agrupamentos de incidentes criminais. Esses centroides, destacados em vermelho na figura, foram sugeridos como os locais ótimos para posicionamento das viaturas para maximizar a cobertura das áreas de alta criminalidade.

Para maximizar a cobertura e a resposta rápida às áreas com maior ocorrência de crimes, as viaturas policiais devem ser posicionadas próximas aos centroides dos clusters. Isso garante que a presença policial está centralizada em áreas com alta densidade criminal, facilitando a prevenção e a resposta rápida a incidentes.

Em síntese, este capítulo demonstra que técnicas de *machine learning*, quando integradas com análises espaciais e temporais, podem proporcionar *insights* valiosos para a predição criminal. A abordagem adotada oferece uma base sólida para o desenvolvimento de modelos preditivos que podem auxiliar na prevenção e na gestão de crimes, promovendo uma maior segurança nas comunidades analisadas.

DADOS CURRICULARES

IDENTIFICAÇÃO	
	ANDERSON GARRIDO SCAIONI 01/10/1977
Nacionalidade	brasileira
Nome em citações bibliográficas:	Scaioni, Anderson Garrido Scaioni, A.G.
Currículo Lattes	http://lattes.cnpq.br/8882314092408579
ORCID	https://orcid.org/0009-0003-2263-5218
FORMAÇÃO ACADÊMICA	
2022/2025	Mestrado em Ciência da Computação. Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, Brasil. Título: Policiamento Inteligente: Aplicação de Ontologia, Análise Preditiva e Redes Neurais Convolucionais no Apoio ao Planejamento Operacional e Roteirização. Orientador: Rogério Eduardo Garcia. Setores de atividade: Administração pública, defesa e seguridade social.
2012/2014	Graduação em Bacharelado em ciência policiais e ordem pública. Academia de Polícia Militar do Barro Branco, APMBB, Brasil. Título: Tecnologias de Leitura Óptica utilizadas na PMESP. Orientador: Michelina Toniato. Bolsista do(a): Polícia Militar do Estado de São Paulo, PMESP, Brasil.
2007/2010	Graduação em Tecnologia em Desenvolvimento Web. Universidade do Oeste Paulista, UNOESTE, Brasil. Título: Serviço de Dia On_Line - SDOL. Orientador: Eliezer Gomes Parangaba Filho. Bolsista do(a): Universidade do Oeste Paulista, UNOESTE, Brasil.
1996/2000	Graduação em Direito. Universidade do Oeste Paulista, UNOESTE, Brasil. Título: Usucapião de Veículos de Procedência Criminosa. Orientador: Orlando Monteiro do Amaral Junior.

PATENTES E REGISTROS (Programa de Computador)

SCAIONI, A. G.. SISTEMA ÓRION - BASE INFORMATIZADA DE PREVENÇÃO CRIMINAL. 2016.
Patente: Programa de Computador. Número do registro: 2016000395-2, data de registro: 16/08/2016, título: "SISTEMA ÓRION - BASE INFORMATIZADA DE PREVENÇÃO CRIMINAL" , Instituição de registro: INPI - Instituto Nacional da Propriedade Industrial.

PARTICIPAÇÃO EM EVENTOS CIENTÍFICOS

COWORKINK FOMENTA VALE,27., 2023, (Assis/SP). Mini-curso: Análise Comportamental na gestão de TI. apresentado: Abordagens para entender e adaptar a experiência do usuário em Projetos de Gestão de TI no contexto de roteirização inteligente. 2023. (Seminário).

COWORKINK FOMENTA VALE, 27., 2023, (Assis/SP). Mini-curso: Análise Preditiva em Sprints. apresentado: Tomadas de decisões rápidas e informadas no desenvolvimento Scrum. 2023. (Seminário).

COWORKINK FOMENTA VALE, 28., 2023, (Assis/SP). Mini-curso: Integração de desenvolvimento Ágil Scrum. apresentado: Aplicando Redes Neurais Convolucionais em projetos de TI. 2023. (Seminário).

COWORKINK FOMENTA VALE, 28., 2023, (Assis/SP). Mini-curso: Ontologias dinâmicas no ciclo de desenvolvimento Scrum. apresentado: Estratégias para melhorar a compreensão e colaboração. 2023. (Seminário).