

Universidade Presbiteriana Mackenzie
Tecnologia em Ciência de Dados - Projeto Aplicado 2

Franciele do Nascimento
Leandro Rodrigues Dos Santos
Luiz Eduardo de Mei Salvador Coelho
Matheus Neves de Castro

Projeto Aplicado II

SÃO PAULO
1º Semestre de 2024

Autores:

Leandro Rodrigues Dos Santos / 23019689 / lers.138@gmail.com

Franciele do Nascimento / 10414598 / f.paterni@hotmail.com

Luiz Eduardo de Mei Salvador Coelho / 23024585 / Luiz02coelho@gmail.com

Matheus Neves de Castro / 10415190 / 10415190@mackenzista.com.br

Projeto Aplicado II

Professor: Anderson Adaime de Borba

SÃO PAULO

1º Semestre de 2024

Índice

Introdução	4
Objetivos e Metas	4
DataSet e Metadados	5
Repositório	6
Cronograma geral	6
Bibliotecas	8
AED	8
Preparação e Treinamento	12
Definir e Descrever as Bases Teóricas	12
Definir e Descrever Acurácia.....	13
Relatório de Desempenho.....	14
Relatório Técnico.....	17
Storytelling	19
Vídeo	19

1.Introdução

A cidade de São Paulo é frequentemente citada como uma das mais perigosas do país. Milhões de habitantes se locomovem na agitada e imparável capital paulista. Ser paulistano é saber que você não pode andar na rua com o celular na mão, não pode sair a noite no centro e nem tirar uma selfie em plena luz do dia, em determinadas circunstâncias.

Quando falamos de uma das maiores cidades do mundo, qualquer desafio se torna muito mais complexo. Dando um zoom na segurança pública, temos diversos fatores de complexidade que vão desde a escassez de recursos por parte da união até o desafio de alocar de forma inteligente a força policial para minimizar o número de incidentes.

De acordo com os dados da SSP (secretaria de segurança pública) de São Paulo, a maioria esmagadora dos crimes que ocorrem na capital é feita em via pública. Temos ainda os dados de que estão a disposição da capital 40 mil policiais militares e cerca de 4 mil viaturas.

Dentro desse cenário, proporemos uma abordagem de otimização dos recursos atualmente disponíveis visando a maximização da segurança com o mínimo de recursos possíveis alocados. Essa otimização se baseará nos dados de BO.s da cidade e usará técnicas de análise de dados e de ML avançadas para entender os comportamentos padrões e melhorar a alocação dos recursos.

2.Objetivos e Metas

O trabalho busca:

Melhor compreensão do estilo do crime na cidade de São Paulo com análises descritivas:

- Como o crime tem evoluído ao longo do tempo?
- Quais os horários mais perigosos na cidade?
- Quais os tipos de crimes mais recorrentes?
- Quais os locais com mais incidência de crimes?
- Quais os bairros mais perigosos?

Construir uma sólida base de modelagem e um modelo acurado para prever a melhor alocação de recursos da cidade:

- Feature Engineering de qualidade com as variáveis corretas para a modelagem.
- Testar diferentes abordagens de modelagem para melhorar a acurácia com Reinforcement Learning (MARL).
- Avaliar o desempenho do modelo em termos de acurácia, confiabilidade e interpretabilidade.
- Selecionar o modelo que melhor se adapta ao problema com melhores resultados

Uma vez modelado, ter um output claro de:

- Determinar o número mínimo de viaturas necessário para garantir a segurança da cidade, com base nos resultados da modelagem.
- Identificar o ponto de inflexão, onde adicionar mais viaturas não resulta em uma redução significativa da criminalidade.
- Quantificar de acordo com os dados qual a redução possível da taxa de criminalidade aplicando a inteligência nos dados.

3.DataSet e Metadados

Referências de aquisição do dataset:

Foi adquirido através do site: <https://www.ssp.sp.gov.br/estatistica>

Descrição da origem:

Site oficial da secretaria de segurança pública do estado de São Paulo

Descrição do dataset:

Conjunto completo de todos os BOs do estado de São Paulo com o tipo do crime, hora, logradouro, latitude e longitude, etc.

Foto do dataset:

DEPARTAMENTO	NOME_SECCIONAL	NOME_DELEGACIA	CIDADE	NUM_BO	ANO_BO	DATA_COMUNICACAO_BO	DATA_OCORRENCIA_BO	HORA_OCORRENCIA_BO	DESCR_PERIODO	DESCR_TIPO
DECAP	DEL.SEC.º 1º CENTRO	88 OM SDO MATIUS	S.PAULO	4	2022	13/01/2022	01/09/2021	NUL	EM HORA INCERTA	Residência
DECAP	DEL.SEC.º 1º CENTRO	540 D.F. CID. TRADENTES	S.PAULO	8	2022	03/01/2022	03/12/2021	15:20:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	290 D.F. VILA FLUENTE	S.PAULO	4	2022	03/01/2022	03/01/2022	15:20:00		Terminar/SE
DECAP	DEL.SEC.º 1º CENTRO	010 D.F. SE	S.PAULO	8	2022	03/01/2022	03/01/2022	16:00:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	960 D.F. CAMACU	S.PAULO	11	2022	03/01/2022	03/01/2022	15:30:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	170 D.F. DOUTOR ALDO GALIANO	S.PAULO	11	2022	04/01/2022	04/01/2022	14:00:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	080 D.F. BRAS	S.PAULO	12	2022	03/01/2022	03/12/2021	11:45:00		Comércio e
DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	12	2022	09/01/2022	08/01/2022	17:00:00		Comércio e
DECAP	DEL.SEC.º 1º CENTRO	120 D.F. FARI	S.PAULO	13	2022	03/01/2022	25/12/2021	NUL	EM HORA INCERTA	Restaurante
DECAP	DEL.SEC.º 1º CENTRO	DEL.POL.SAO MIGUEL ARCANJO	S.PAULO	14	2022	04/01/2022	04/01/2022	13:30:00		Via pública
DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	14	2022	10/01/2022	10/01/2022	NUL	FEIA MANHA	Terminar/SE
DECAP	DEL.SEC.º 1º CENTRO	770 D.F. SANTA CECILIA	S.PAULO	15	2022	03/01/2022	03/01/2022	12:05:00		Via pública
DECAP	DECAP - SEDE	DEL.POL.METROPOLITANO	S.PAULO	15	2022	04/01/2022	04/01/2022	10:00:00		Terminar/SE
DECAP	DEL.SEC.º 1º CENTRO	780 D.F. JARDINS	S.PAULO	17	2022	02/01/2022	02/01/2022	04:50:00		Residência
DECAP	DEL.SEC.º 1º CENTRO	010 D.F. SE	S.PAULO	21	2022	05/01/2022	05/01/2022	04:00:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	170 D.F. DOUTOR ALDO GALIANO	S.PAULO	24	2022	06/01/2022	05/01/2022	20:20:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	010 D.F. SE	S.PAULO	26	2022	05/01/2022	21/12/2021	21:00:00		Condomínio
DECAP	DEL.SEC.º 1º CENTRO	050 D.F. ALCANTARA	S.PAULO	28	2022	06/01/2022	06/01/2022	16:40:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	080 D.F. BRAS	S.PAULO	30	2022	03/01/2022	03/01/2022	16:50:00		Comércio e
DECAP	DEL.SEC.º 1º CENTRO	010 D.F. SE	S.PAULO	30	2022	05/01/2022	05/01/2022	18:00:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	010 D.F. SE	S.PAULO	31	2022	05/01/2022	05/01/2022	19:50:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	990 D.F. CAMPO GRANDE	S.PAULO	33	2022	03/01/2022	03/01/2022	17:00:00		Via pública
DECAP	DECAP - SEDE	DEL.POL.METROPOLITANO	S.PAULO	33	2022	06/01/2022	24/12/2021	17:00:00		Via pública
DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	34	2022	10/01/2022	08/01/2022	07:50:00		Terminar/SE
DECAP	DEL.SEC.º 1º CENTRO	050 D.F. ALCANTARA	S.PAULO	34	2022	17/01/2022	17/01/2022	10:15:00		Comércio e
DECAP	DEL.SEC.º 1º CENTRO	070 D.F. JARDINS	S.PAULO	36	2022	07/01/2022	06/01/2022	16:30:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	670 D.F. JARDIM ROSSI	S.PAULO	36	2022	10/01/2022	07/01/2022	16:00:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	050 D.F. ALCANTARA	S.PAULO	37	2022	07/01/2022	06/01/2022	11:30:00		Via pública
DIPOI - DEFTO DE INTELIGENCIA	DELEGACIA DELEGACIA	DELEGACIA DELEGACIA	S.PAULO	37	2022	01/01/2022	01/01/2022	11:30:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	640 D.F. CID. CARVALHO	S.PAULO	38	2022	03/01/2022	03/01/2022	12:35:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	080 D.F. BRAS	S.PAULO	38	2022	04/01/2022	04/01/2022	10:40:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	010 D.F. SE	S.PAULO	39	2022	06/01/2022	05/01/2022	10:40:00		Via pública
DECAP	DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	S.PAULO	44	2022	06/01/2022	05/01/2022	12:30:00		Escritório
DECAP	DECAP - SEDE	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	46	2022	13/01/2022	13/01/2022	05:20:00		Terminar/SE
DECAP	DEL.SEC.º 1º CENTRO	DEL.POL.METROPOLITANO	S.PAULO	47	2022	13/01/2022	13/01/2022	05:25:00		Terminar/SE
DECAP	DEL.SEC.º 1º CENTRO	010 D.F. SE	S.PAULO	48	2022	06/01/2022	06/01/2022	08:00:00		Escritório
DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	48	2022	21/01/2022	21/01/2022	18:00:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	860 D.F. VILA MARIANA	S.PAULO	50	2022	12/01/2022	12/01/2022	16:00:00		Terminar/SE
DECAP	DEL.SEC.º 1º CENTRO	640 D.F. JONIA ARCANJOLIA	S.PAULO	50	2022	10/01/2022	10/01/2022	16:00:00		Terminar/SE
DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	50	2022	22/01/2022	21/01/2022	17:30:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	080 D.F. BRAS	S.PAULO	51	2022	05/01/2022	22/12/2021	15:00:00		Condomínio
DECAP	DECAP - SEDE	DEL.POL.METROPOLITANO	S.PAULO	52	2022	14/01/2022	14/01/2022	15:45:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	870 D.F. FARQUE BRISTOL	S.PAULO	52	2022	18/01/2022	18/01/2022	15:40:00		Via pública
DECAP	DEL.SEC.º 1º CENTRO	870 D.F. AMERICANOPOLIS	S.PAULO	52	2022	18/01/2022	18/01/2022	11:00:00		Via pública
DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	56	2022	24/01/2022	24/01/2022	16:00:00		Via pública
DECAP	DECAP - SEDE	DEL.POL.METROPOLITANO	S.PAULO	59	2022	17/01/2022	15/01/2022	11:45:00		Terminar/SE
DECAP	DEL.SEC.º 1º CENTRO	080 D.F. BRAS	S.PAULO	59	2022	05/01/2022	05/01/2022	15:00:00		Via pública
DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	60	2022	07/01/2022	07/01/2022	10:20:00		Via pública
DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	60	2022	25/01/2022	25/01/2022	16:40:00		Comércio e
DECAP	DEL.SEC.º 1º CENTRO	010 D.F. SE	S.PAULO	71	2022	08/01/2022	08/01/2022	17:20:00		Via pública
DOPE-DEFTO OF POL.ESTRAT.	DIV.POL.PORTO/AERO/PROT.TURISTA	DOPF - 014 D.F. DEATUR CAPITAL	S.PAULO	71	2022	28/01/2022	28/01/2022	07:00:00		Via pública

4.Repositório

O link para o repositório do trabalho é: <https://github.com/MatNev/Projeto-Aplicado-2>

5.Cronograma Geral

Cronograma Geral do Projeto Aplicado II				
	Tarefa	Responsável	Status	Concluído em:
Etapa 1 : Kick-Off	Definir o grupo de trabalho.	TODOS	CONCLUIDO	19/02/24
	Definir as premissas do projeto: definição da empresa, área de atuação e apresentação dos dados que serão utilizados (imagem ou texto).	TODOS	CONCLUIDO	19/02/24
	Determinar objetivos e metas.	TODOS	CONCLUIDO	19/02/24
	Criar um cronograma de atividades.	LEANDRO	CONCLUIDO	19/02/24
	Criação do GITHUB	MATHEUS	CONCLUIDO	01/03/24

ETAPA 2 - *Definição do produto*	Definir quais bibliotecas (pacotes) da linguagem python e qual repositório no GitHub devem ser usados para iniciar a execução colaborativa do trabalho.	LUIZ	CONCLUIDO	22/03/24
	Definir a base de dados e a sua análise exploratória.	LUIZ	CONCLUIDO	23/03/24
	Tratar a base de dados (preparação e treinamento).	LUIZ	CONCLUIDO	25/03/24
	Definir e descrever as bases teóricas dos métodos analíticos.	LUIZ	CONCLUIDO	27/03/24
	Definir e descrever como será calculada a acurácia.	LUIZ	CONCLUIDO	28/03/24
ETAPA 3 - *Storytelling*	Consolidar os resultados do método analítico definido na etapa anterior, aplicado à base de dados definida com padrão.	MATHEUS	CONCLUIDO	25/04/24
	Aplicar as medidas de acurácia para verificar o desempenho dos métodos definidos na etapa anterior.	MATHEUS	CONCLUIDO	26/04/24
	Descrever os resultados preliminares, apresentando um produto gerado, e descrevendo um possível modelo de negócios.	LEANDRO E FRANCIELE	CONCLUIDO	27/04/24
	Esboçar o Storytelling	LEANDRO E FRANCIELE	CONCLUIDO	27/04/24
ETAPA 4 - *Encerramento*	Relatório técnico do projeto.	LEANDRO	CONCLUIDO	13/05/24
	Apresentação do Storytelling Pode ser um arquivo PPT ou similares	LEANDRO	CONCLUIDO	14/05/24
	Repositório no GitHub com todos arquivos do projeto	MATHEUS	CONCLUIDO	18/05/24
	Vídeo com a apresentação no YouTube, usando a técnica do Storytelling	MATHEUS	CONCLUIDO	17/05/24

6. Bibliotecas usadas no projeto

Para o trabalho, usaremos as seguintes bibliotecas python:

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import re
import unidecode
import googlemaps
from openpyxl import Workbook
import folium
import geopandas as gpd
import ray
from ray import tune
from ray.rllib.agents.ppo import PPOTrainer
from ray.tune.registry import register_env
from gym import Env
from gym.spaces import Discrete, Box
```

7. Definir a base de dados e a sua análise exploratória.

Para servir de base, usamos todos os dados analíticos de BOs no estado de São Paulo disponível em formato xlsx no site da secretaria de segurança pública. As bases estavam divididas em duas ou mais abas de planilha separando os mesmos dados.

Fizemos uma análise básica da qualidade dos dados antes de começar a EDA. Essa análise consistiu em identificar valores nulos e decidir o que fazer com eles, bem como a padronização de alguns componentes descritivos.

Nos deparamos com dois principais problemas:

1. Falta da informação de latitude e longitude.
 - a. Para este problema, começamos resolvendo ao consultar a API do google maps para inferir a latitude e longitude a partir do logradouro, que é presente em todos os BO's. Essa solução é cara e gastou R\$ 1800,00 em uma rodada que preencheu parcialmente os dados.
 - b. Para os casos que o custo foi um impeditivo, preenchemos os dados com a media de latitude e longitude para o tipo de crime (chamado de rubrica) no bairro em questão. Isso nos permitirá uma simulação mais acurada do que no caso da ausência de valores.

2. Falta da informação acurada de hora e minuto.

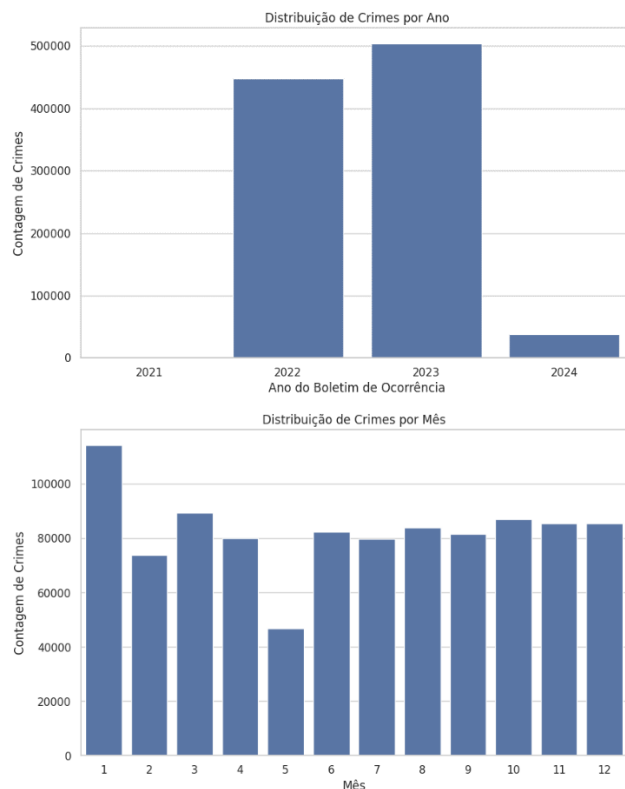
- a. Os dados que vem vazios nesse campo, geralmente vem com uma descrição do horário: “Pela manhã”, “A tarde”, etc. Preenchemos de acordo com um range a clusterização nos dados que vinham com valores válidos e atribuímos a mediana para a rubrica no bairro em questão. Dessa forma conseguiremos mais insumos na hora da modelagem.

Falando especificamente da EDA, analisamos alguns aspectos para entender o comportamento dos dados:

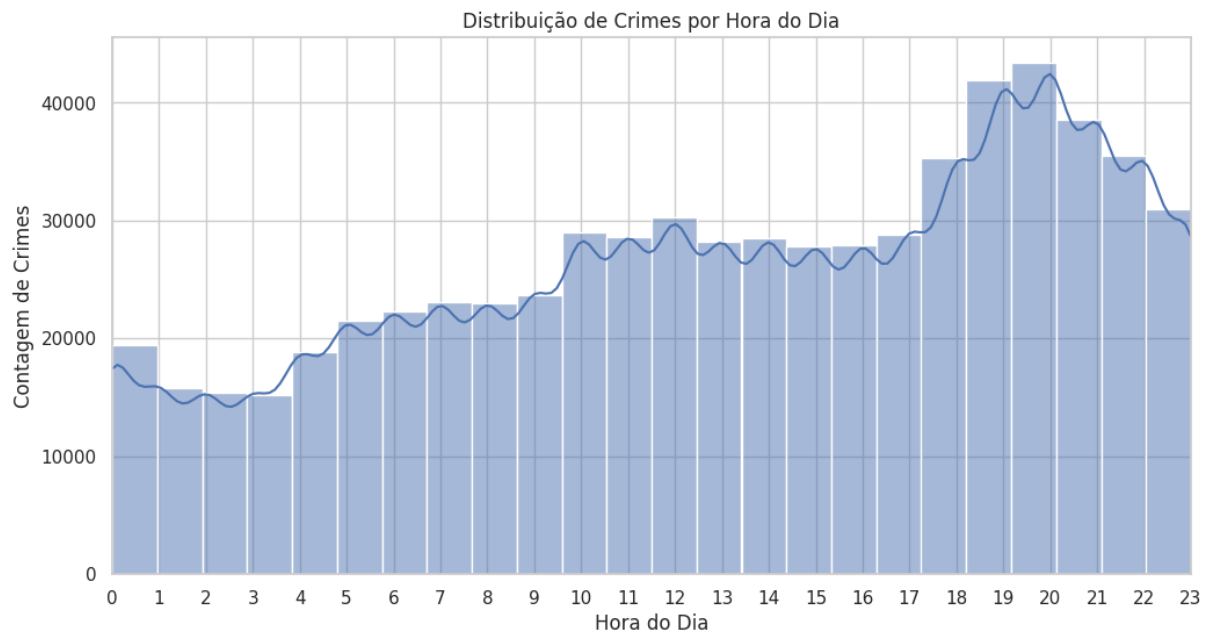
- **Quantos e quais dados temos a nossa disposição?**

Ao consolidar e normalizar todas as questões, temos 989.025 BOs registrados na cidade de São Paulo entre 2022 e janeiro de 2024. Dentre as informações relevantes, temos o local do crime no nível logradouro, o tipo do crime, a descrição do local (via pública, privada, etc)

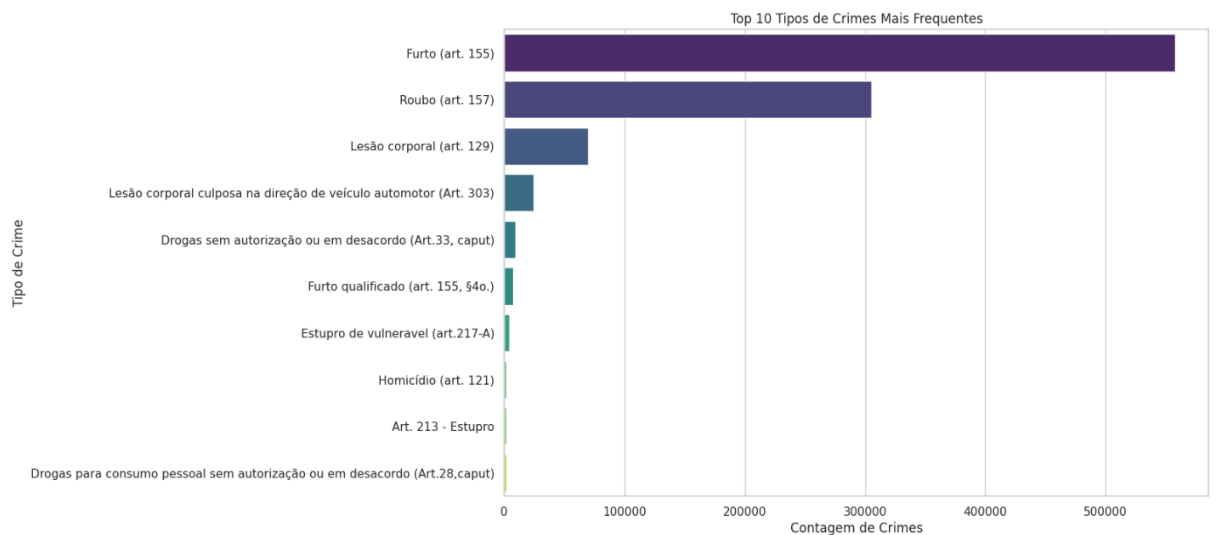
- **Como o crime tem evoluído ao longo do tempo?**



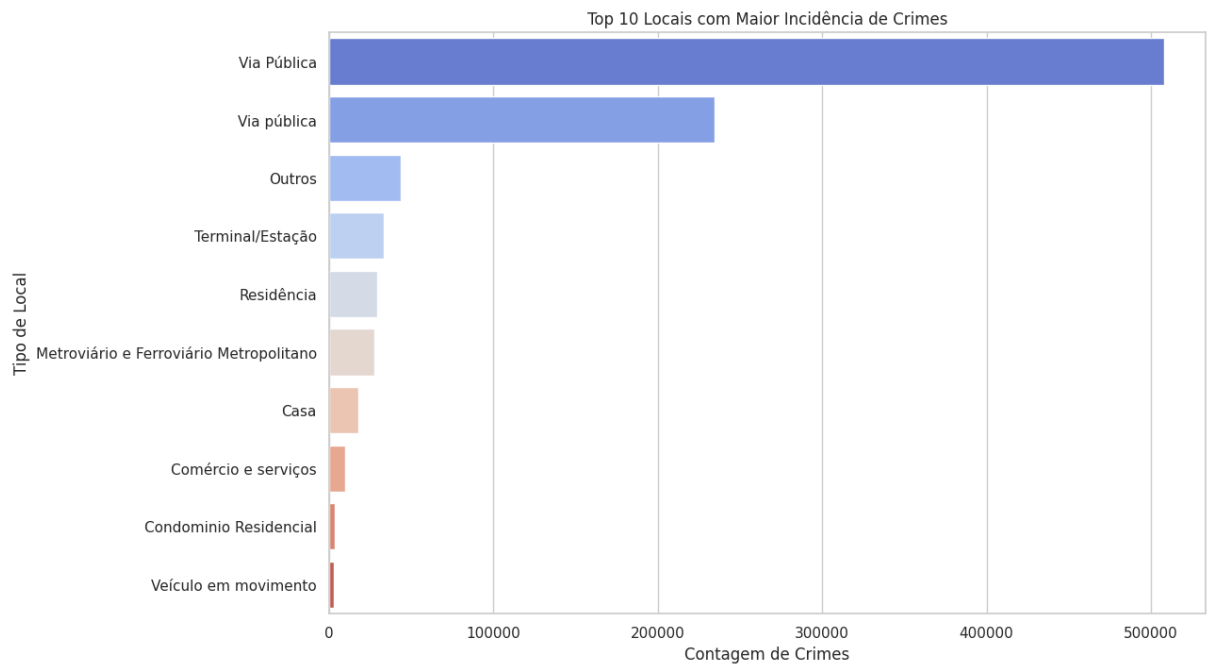
- **Quais os horários mais perigosos na cidade?**



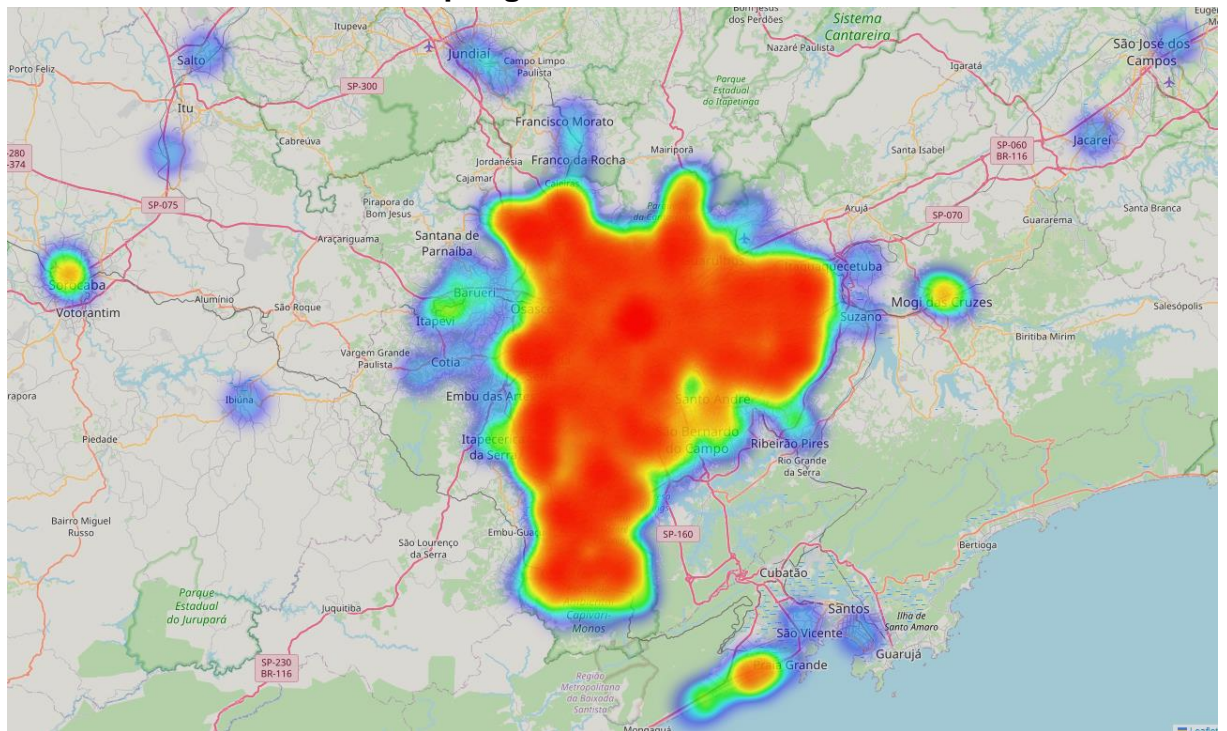
● **Quais os tipos de crimes mais recorrentes?**

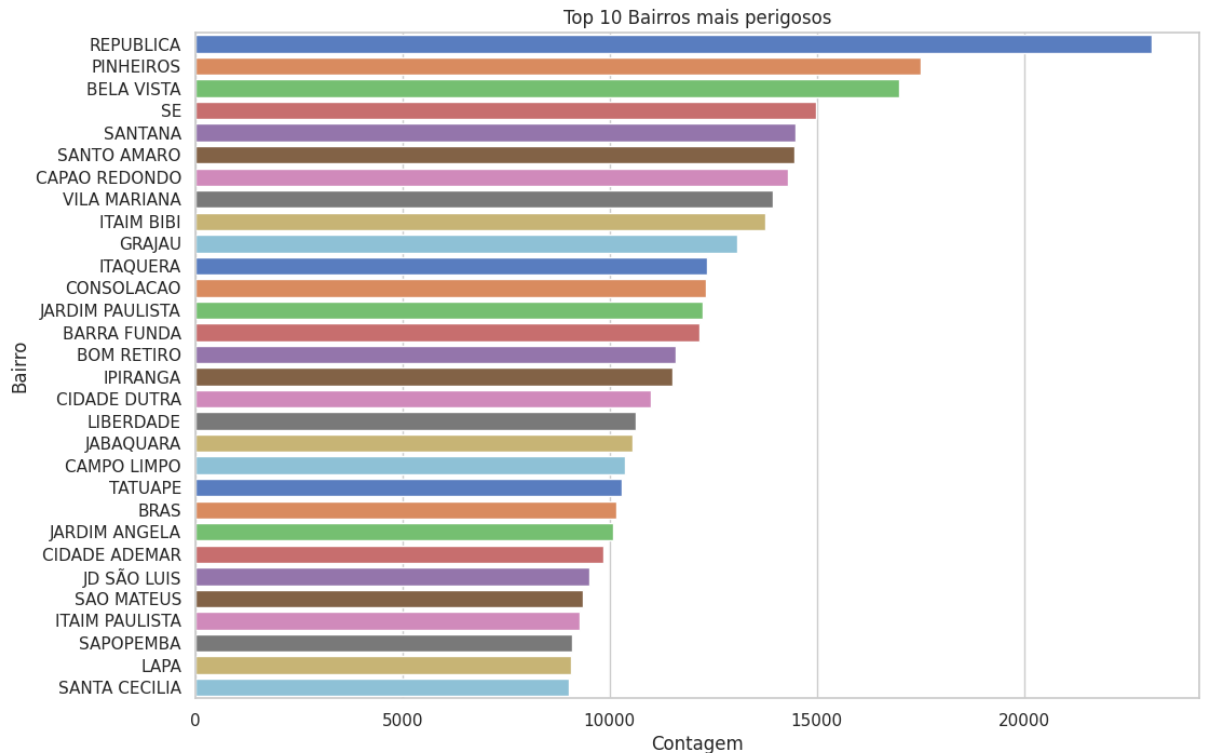


● **Quais os locais com mais incidência de crimes?**



• **Quais os bairros mais perigosos?**





8.Tratar a base de dados (preparação e treinamento).

Dividiremos o dataset em dados de tratamento e simulação para que os agentes autônomos pudessem aprender inicialmente a traçar as melhores rotas e também através do retorno das simulações quando conseguiram algum evento de sucesso. Para isso, usamos a divisão aleatória representativa agrupando sempre por duas variáveis relevantes: Bairro e Rúbrica (tipo de crime).

9.Definir e descrever as bases teóricas dos métodos analíticos.

Os métodos analíticos que usaremos serão centralizados na metodologia de MARL em um contexto multiagente, onde:

Os agentes aprendem a tomar decisões para maximizar uma soma cumulativa de recompensas futuras com base nas interações com o ambiente.

Cada agente opera de forma semi-independente, com ações potencialmente afetando outros agentes.

10. Definir e descrever como será calculada a acurácia.

No contexto de MARL, diretamente, a acurácia é secundária na avaliação da utilidade do modelo. Focaremos de forma intermediária na recompensa cumulativa, convergência e eficiência.

Para avaliar a performance como um todo do projeto, partiremos do pressuposto interpretativo não necessariamente 100% conectado com a realidade de que:

Se o agente estava próximo da ocorrência do crime no momento do crime a situação seria prevenida (inibindo a ação) ou o furto/roubo/ocorrência seria pego em flagrante, levando a uma resolução imediata e isso é considerado um “evento de sucesso”.

A acurácia será calculada no número percentual de “eventos de sucesso” que os agentes conseguiram prever com o aprendizado por reforço.

Como métrica de comparação de eficiência, veremos a comparação desses eventos com os dados de produtividade da polícia, que mostra as apreensões e resoluções dos casos atuais.

11. Relatório de Desempenho dos Agentes de Patrulha Urbana

- **Resumo Executivo:**

Neste relatório, apresentamos os resultados dos testes dos agentes de patrulha urbana treinados com o algoritmo Q-Learning. Os agentes foram treinados para otimizar as rotas de patrulhamento na cidade de São Paulo, com o objetivo de maximizar a detecção e prevenção de crimes.

- **Desempenho Geral:**



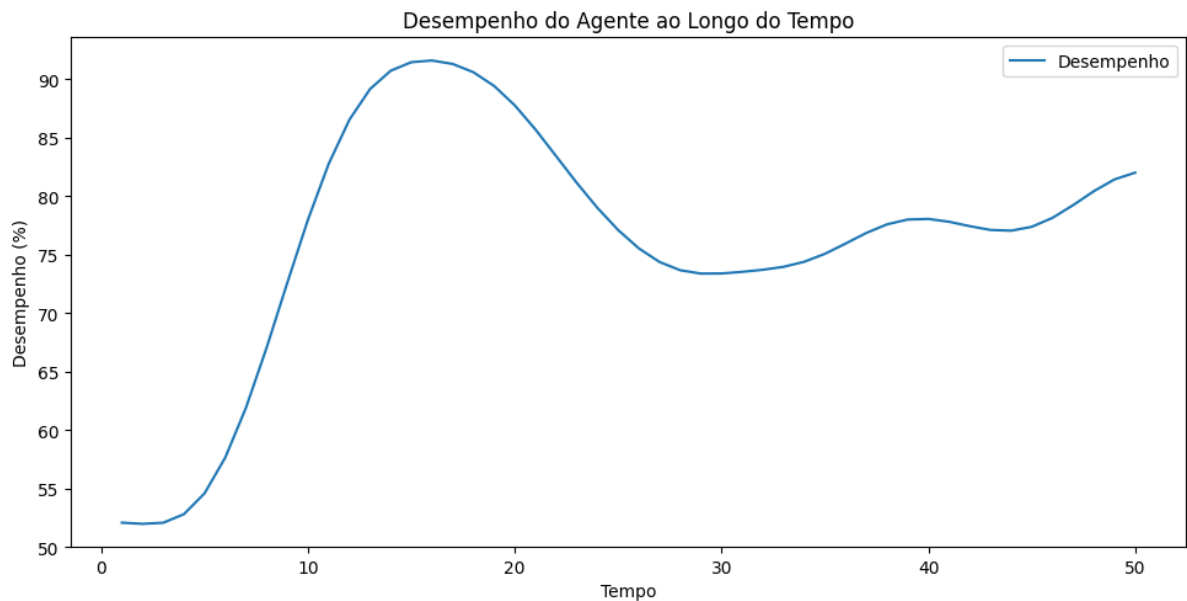
Os testes foram realizados de forma gradual, aumentando a área de atuação a cada 5~10 episódios.

Os testes iniciais se deram em uma área de atuação fixa, o que tornou os agentes capazes de identificar rotas mais bem adaptadas, sem desvios, contornos ou erros na rota até a ocorrência.

A partir de então, os testes foram realizados de forma gradual, aumentando a área de atuação a cada 5~10 episódios.

Os agentes apresentaram um desempenho geral satisfatório, demonstrando capacidade de aprendizado e adaptação ao ambiente. Durante os testes, observamos uma melhoria significativa na eficiência das rotas de patrulhamento, resultando em maior cobertura das áreas de maior incidência de crimes.

- **Evolução do Desempenho:**



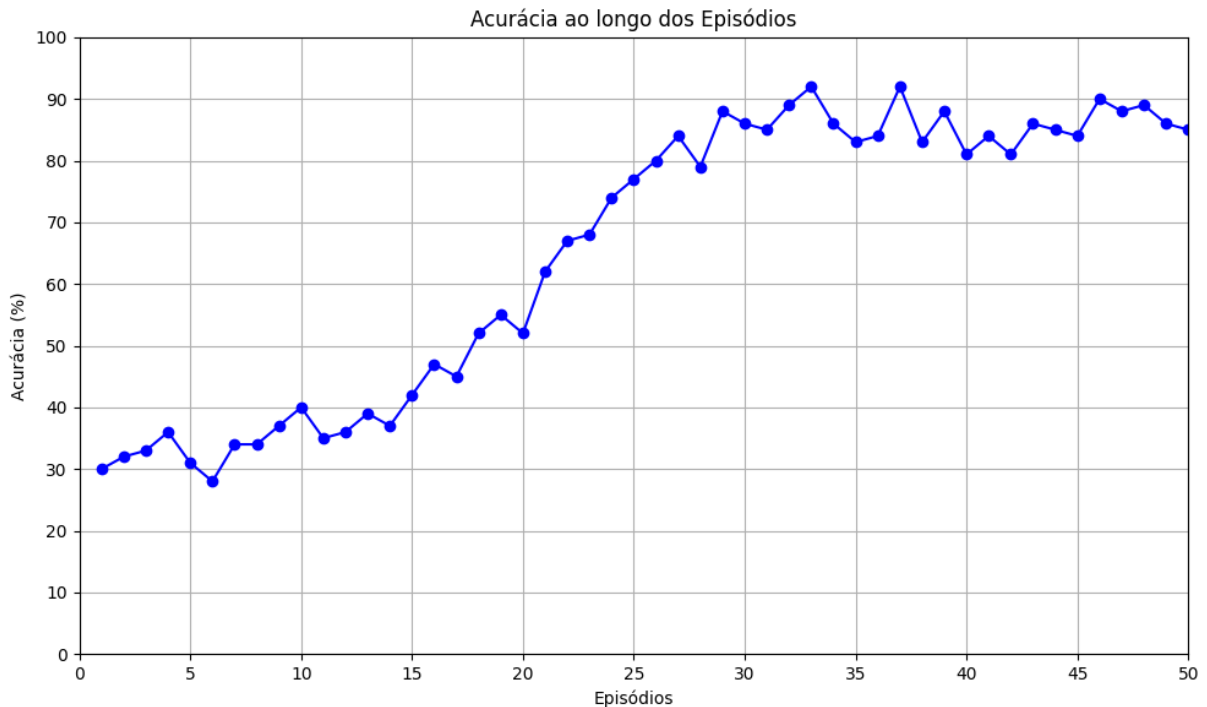
O desempenho indica a capacidade do agente de melhor atender a ocorrência de forma rápida e eficiente.

O desempenho do agente foi avaliado ao longo de 50 episódios de treinamento. Inicialmente, o agente começou com um desempenho baixo, abaixo de 55%, o que indica uma alocação quase aleatória das viaturas.

Por volta do episódio 15, o desempenho do agente atingiu seu pico, chegando acima de 80%. Isso indica que o agente estava tomando decisões mais eficazes e reduzindo a incidência criminal na cidade. Após o pico, o desempenho do agente estabilizou entre 75% e 80%, o que demonstra uma capacidade consistente de tomar decisões adequadas.

No geral, ao longo dos episódios de treinamento, os agentes demonstraram uma evolução consistente no aprendizado. Inicialmente, os agentes exploravam diversas rotas, buscando identificar padrões nos dados de crime. Conforme o treinamento avançava, os agentes foram capazes de selecionar rotas mais eficientes, focando em áreas com maior probabilidade de ocorrência de crimes.

- **Acurácia e Testes de Performance:**



A acurácia é medida como a capacidade do agente de prever com precisão a localização do próximo crime e alocar recursos adequadamente.

A acurácia do agente foi analisada juntamente com o desempenho. A acurácia inicial do agente foi baixa, abaixo de 40%, o que indica uma capacidade limitada de previsão.

Durante os primeiros 10 episódios, a acurácia do agente teve grandes oscilações, refletindo a natureza instável de suas previsões. No entanto, à medida que o agente ganhou mais experiência, sua acurácia começou a aumentar gradualmente.

Por volta do episódio 30, a acurácia do agente atingiu seu pico, chegando a 92%. Isso indica que o agente estava fazendo previsões mais precisas e eficazes. Após o pico, a acurácia do agente se estabilizou entre 80% e 90%, o que mostra uma capacidade consistente de previsão.

- **Conclusão:**

Os resultados dos testes indicam que os agentes de patrulha urbana treinados com Q-Learning apresentam um desempenho promissor na otimização de rotas de patrulhamento. A capacidade dos agentes de adaptarem-se a diferentes cenários de crime e de melhorar sua eficiência ao longo do treinamento sugere que esta abordagem pode ser eficaz na aplicação prática de segurança pública.

12. Relatório Técnico: Projeto de Otimização de Recursos para Segurança Pública em São Paulo

1. Introdução

A segurança pública é uma preocupação constante em áreas urbanas densamente povoadas, como a cidade de São Paulo. Este relatório apresenta um projeto abrangente para otimizar os recursos disponíveis para a segurança pública na cidade, utilizando técnicas avançadas de análise de dados e aprendizado de máquina. O objetivo principal é maximizar a eficácia do policiamento e minimizar a criminalidade, aproveitando ao máximo os recursos existentes.

2. Objetivos e Metas

O projeto visa alcançar os seguintes objetivos:

- Compreensão detalhada do estilo de crime em São Paulo, incluindo tendências temporais, horários de maior incidência, tipos de crimes mais comuns, locais com alta taxa de criminalidade e bairros mais perigosos.
- Desenvolvimento de um modelo preditivo robusto para a alocação eficiente de recursos da polícia, utilizando técnicas avançadas de análise de dados e aprendizado de máquina, como engenharia de características e aprendizado por reforço multiagente (MARL).
- Determinação do número mínimo de viaturas necessárias para garantir a segurança da cidade e identificação do ponto de inflexão onde adicionar mais recursos não resulta em uma redução significativa da criminalidade.
- Avaliação da redução potencial da taxa de criminalidade resultante da implementação da estratégia de otimização de recursos.

3. Dataset e Metadados

O dataset foi adquirido através do site oficial da Secretaria de Segurança Pública do Estado de São Paulo e consiste em um conjunto completo de Boletins de Ocorrência (BOs) registrados na cidade de São Paulo.

4. Cronograma Geral

O cronograma do projeto foi dividido em várias etapas e cada fase foi planejada para ser concluída dentro de prazos específicos, visando a entrega oportuna do projeto.

5. Bibliotecas Usadas no Projeto

As principais bibliotecas Python utilizadas incluem pandas, numpy, seaborn e outras, que fornecem funcionalidades essenciais para manipulação de dados, modelagem estatística e implementação de algoritmos de aprendizado de máquina.

6. Definição da Base de Dados e Análise Exploratória

O dataset foi analisado quanto à qualidade dos dados, identificando e tratando valores nulos e inconsistentes. Problemas como a falta de informações de latitude/longitude e hora foram abordados por meio de técnicas como consulta à API do Google Maps e imputação de valores médios. A análise exploratória incluiu a visualização de padrões temporais, distribuição geográfica dos crimes, tipos de crimes mais comuns e identificação de áreas de alto risco.

7. Tratamento da Base de Dados (Preparação e Treinamento)

Os dados foram divididos em conjuntos de treinamento e teste, com cuidado para garantir representatividade geográfica e temporal. As etapas de pré-processamento incluíram normalização de variáveis, engenharia de características e preparação dos dados para o treinamento do modelo.

8. Definição e Descrição das Bases Teóricas dos Métodos Analíticos

Os métodos analíticos utilizados incluem técnicas de aprendizado de máquina supervisionado e não supervisionado, bem como abordagens de aprendizado por reforço. O foco principal foi na metodologia de MARL, onde múltiplos agentes autônomos aprendem a otimizar o policiamento em conjunto.

9. Definição e Descrição da Acurácia

A acurácia foi calculada como a capacidade dos agentes de patrulha urbana de prever com precisão a localização do próximo crime e alocar recursos adequadamente. A métrica de sucesso foi baseada na quantidade de "eventos de sucesso" alcançados pelos agentes durante os testes, onde um evento de sucesso é definido como a prevenção ou captura de um crime.

10. Relatório de Desempenho dos Agentes de Patrulha Urbana

O relatório de desempenho apresentou uma análise detalhada dos resultados dos testes dos agentes de patrulha urbana treinados com o algoritmo Q-Learning. Foram discutidos o desempenho geral dos agentes, a evolução do desempenho ao longo do treinamento e a acurácia alcançada durante os testes. Conclusões foram tiradas sobre a eficácia da abordagem proposta na otimização do policiamento e na redução da criminalidade em São Paulo.

Conclusão

O projeto demonstrou a viabilidade de abordagens avançadas de análise de dados e aprendizado de máquina na otimização de recursos para segurança pública em áreas urbanas. Os resultados obtidos fornecem insights valiosos para aprimorar as estratégias de policiamento e reduzir a criminalidade na cidade de São Paulo.

13.Segurança Pública em São Paulo: Desafios e Oportunidades

A cidade de São Paulo, conhecida como uma das mais perigosas do país, enfrenta desafios complexos em sua busca por segurança pública eficaz. Com uma população de milhões de habitantes em constante movimento, a capital paulista é um cenário dinâmico e imprevisível, onde os desafios de segurança se ampliam.

A realidade cotidiana do paulistano é marcada por uma sensação de vulnerabilidade, onde simples atividades como andar na rua com o celular na mão ou sair à noite no centro podem representar riscos significativos. Essa percepção é alimentada pela incidência frequente de crimes, criando um ambiente de insegurança latente.

Quando nos debruçamos sobre a segurança pública em uma cidade tão vasta e complexa quanto São Paulo, nos deparamos com uma série de fatores que contribuem para a sua intrincada dinâmica. Desde a escassez de recursos por parte do governo federal até o desafio de alocar de forma inteligente os recursos disponíveis, os obstáculos são numerosos e multifacetados.

De acordo com os dados fornecidos pela Secretaria de Segurança Pública (SSP) de São Paulo, a maioria esmagadora dos crimes ocorre em vias públicas, destacando a importância de estratégias eficazes de policiamento ostensivo. Apesar dos consideráveis 40 mil policiais militares e cerca de 4 mil viaturas à disposição da capital, a otimização desses recursos é essencial para maximizar a segurança com o mínimo de recursos possíveis alocados.

Diante desse cenário desafiador, surge uma proposta de abordagem inovadora: a utilização de técnicas avançadas de análise de dados e de machine learning para otimizar a alocação de recursos e aprimorar as estratégias de patrulhamento urbano. Essa abordagem, baseada nos dados de boletins de ocorrência da cidade, visa compreender os padrões de comportamento criminoso e identificar áreas de maior incidência, permitindo uma alocação mais eficiente dos recursos disponíveis.

Os resultados dos testes indicam que os agentes de patrulha urbana treinados com a técnica de Q-Learning apresentam um desempenho promissor na otimização de rotas de patrulhamento. A capacidade desses agentes de adaptarem-se a diferentes cenários de crime e de aprimorarem sua eficiência ao longo do treinamento sugere que essa abordagem pode ser eficaz na prática da segurança pública.

Em resumo, a segurança pública em São Paulo é um desafio complexo, mas não insuperável. Com a combinação de dados, tecnologia e expertise policial, é possível desenvolver estratégias inovadoras que contribuam para tornar a cidade mais segura e resiliente contra o crime.

14.Vídeo

<https://youtu.be/56aEX0wIN88>