

# Improving crime prediction through ensembles

Ruam E. R. C. Pastor, Jair P. de Sales, Adriano F. C. Filho, Paulo S. G. de Mattos Neto

<sup>1</sup>Centro de Informática - Universidade Federal de Pernambuco (CIn/UFPE)  
Caixa Postal 15.064 – 91.501-970 – Recife – Pernambuco – Brasil

{rercp, jps, afcf, psgmn}@cin.ufpe.br

**Abstract.** *Predicting crimes is a challenge due to the complexity of the inherent temporal series in the phenomenon of interest, which often exhibit both linear and nonlinear components. Moreover, due to the dynamic nature of temporal patterns, depending solely on a single model can result in imprecise predictions. In this context, this study proposes the modeling and forecasting of crime temporal series using both individual models and ensembles, encompassing trainable and non-trainable approaches. Furthermore, a combination method is suggested involving the selection of a subset of individual models through a validation step. For the experiment, crime series from different locations, such as Pernambuco, Chicago, and Los Angeles, were employed. Based on the presented results, it is observed that ensemble models exhibited better predictive performance than the individual models used. However, regarding the proposed method, it is concluded that selecting a subset of individual models based on validation could be problematic when the validation sets are not sufficiently representative, whether due to a limited number of observations or distinct characteristics compared to the test set.*

**Resumo.** *Prever crimes é um desafio devido à complexidade das séries temporais inerentes ao fenômeno de interesse, as quais costumemente possuem componentes lineares e não lineares. Além disso, devido à presença de padrões temporais em constante mudança, a utilização de um único modelo pode resultar em previsões imprecisas. Neste contexto, este estudo propõe a modelagem e previsão de séries temporais de crimes utilizando tanto modelos únicos quanto ensembles, incluindo abordagens treináveis e não treináveis. Propõe-se, também, um método de combinação que envolve a seleção de um subconjunto de modelos únicos por meio de uma etapa de validação. Para o experimento, foram utilizadas séries de crimes de diferentes localidades, como Pernambuco, Chicago e Los Angeles. Com base nos resultados apresentados, observa-se que os modelos ensembles apresentaram melhor desempenho preditivo do que os modelos únicos utilizados. No entanto, em relação ao método proposto, conclui-se que a seleção de um subconjunto de modelos únicos com base na validação pode ser problemática quando os conjuntos de validação não são suficientemente representativos, seja devido ao número limitado de observações ou às características distintas em relação ao conjunto de teste.*

## 1. Introdução

Criminalidade é um fenômeno mundial que afeta países em diferentes níveis de desenvolvimento e impactam de maneira direta a economia de um Estado, causando prejuízos

sociais, como a redução da qualidade de vida, a diminuição da mobilidade populacional, a queda dos níveis de turismo e o aumento da desigualdade entre as classes [Detotto and Otranto 2010, Kelly 2000]. Cabe, portanto, ao Estado, a manutenção de uma segurança pública de qualidade.

Uma comunidade socialmente sustentável depende fortemente de minimizar o crime para que as pessoas possam viver de forma pacífica e ativa, enquanto sociedades corruptas não podem prosperar social e economicamente na ausência de paz [Veiga 2014, Kelly 2000]. Conseqüentemente, a análise e previsão de estatísticas criminais é essencial para melhorar a segurança e a proteção da população, auxiliando o poder público a prevenir e combater o crime, mantendo o desenvolvimento sustentável.

Neste contexto, a modelagem e previsão de séries temporais associadas a índices de criminalidade tem sido uma abordagem comumente empregada para avaliar tais fenômenos e auxiliar processos de tomadas de decisão no que concerne à segurança pública (estes trabalhos serão apresentados na Seção 2 - [Safat et al. 2021, Chen et al. 2008, Tariq et al. 2021]. Entretanto, embora modelos preditivos, estatísticos ou de ML, tenham alcançado resultados promissores em outras aplicações [Khairuddin et al. 2019], ainda existem dificuldades quando se trata da previsão de crimes [Sinha and Biswas 2023]. Lidar com dados reais de característica social pode ser extremamente difícil, dado sua complexidade. Portanto, o desenvolvimento de sistemas de previsão precisos para modelagem de séries temporais com esta característica é uma tarefa desafiadora.

Devido à presença de padrões temporais que mudam ao longo do tempo, a adoção de um único modelo pode levar a previsões imprecisas [SILVA et al. 2021]. Neste cenário, a utilização de múltiplos preditores (ensemble), surge como opção frente a utilização de preditores únicos para a previsão [SILVA et al. 2021]. A adoção de um ensemble é promissora porque reduz o risco de selecionar modelos super ou subajustados, equilibra o *trade-off* entre viés e variância e combina os pontos fortes de diferentes modelos, aumentando sua acurácia [Zhang 2007]. Esse trabalho propõe um sistema que tem o objetivo de selecionar e combinar o melhor subconjunto de modelos na tarefa de previsão de séries temporais de crimes. O sistema ainda seleciona a melhor combinação entre estratégias treináveis e não treináveis. Pelo conhecimento dos autores, esse é o primeiro trabalho que aplica ensemble para previsão de séries temporais de crimes, o que contribui com o avanço do estado da arte nesta área.

## **2. Trabalhos relacionados**

Os estudos relacionados à previsão de crimes ganharam força a partir da década de 1980, com o investimento do governo dos Estados Unidos na área. À época, tinham por objetivo, analisar o impacto de diferentes indicadores sociais no cometimento de crimes e evitar o avanço de números que pudessem se tornar alarmantes para o poder público. Deste investimento, surgem alguns trabalhos; entre estes, [Cohen 1981] tinha por objetivo modelar a tendência de crimes a partir de métodos econométricos.

Conforme as pesquisas foram avançando, outros métodos passaram a ser utilizados, como as Redes Neurais Artificiais (RNA) para a realização de previsão e *crime mapping* [Olligschlaeger 1997]. Modelos Autorregressivos Integrados de Médias Móveis (ARIMA) foram também utilizados, neste caso para prever dados de crime contra a pro-

priedade em cidades da China [Safat et al. 2021].

Como os dados relacionados a crimes possuem diferentes características, os trabalhos, hoje, se dão nas mais diversas áreas. McDowall et al. (2011) utilizaram séries temporais longas para a identificação de ciclos sazonais. O resultado aponta que crimes, em sua maioria, seguem ciclos sazonais, mas, que estão padronizados conforme o local ao qual ocorrem [McDowall et al. 2012]. Zhuang et al. (2017) utilizaram modelo recorrente com informações temporais e espaciais para a identificação de *crime hot spots*. Em resumo, tinham como objetivo identificar em quais locais observados existiam maior probabilidade de acontecer eventos criminosos nos passos seguintes da série [Zhuang et al. 2017].

Recentemente, existe um esforço no sentido de gerar modelos capazes de realizar a previsão de séries temporais de maneira minimamente generalizável. Safat et al. (2021), realizaram um estudo comparativo entre modelos estatísticos tradicionais (ARIMA) e modelos de *deep learning* (DL), Long Short-Term Memory (LSTM), buscando os melhores desempenhos em séries relacionadas a crimes de cidades dos Estados Unidos: Los Angeles e Chicago [Safat et al. 2021]. Neste mesmo sentido, Tariq et al. (2021), fizeram comparações entre os modelos, ARIMA, Holt Winters, Simple Exponential Smoothing e RNN/LSTM. Em ambos os casos os modelos foram treinados de maneira individual [Tariq et al. 2021].

Anuvarshini et al. (2022) desenvolveram um trabalho de cunho teórico. Para isto, realizaram testes com modelos de *machine learning* (ML) e DL. Entre os modelos testados, a *Convolutional Neural Network* (CNN) obteve o melhor desempenho [Anuvarshini et al. 2022].

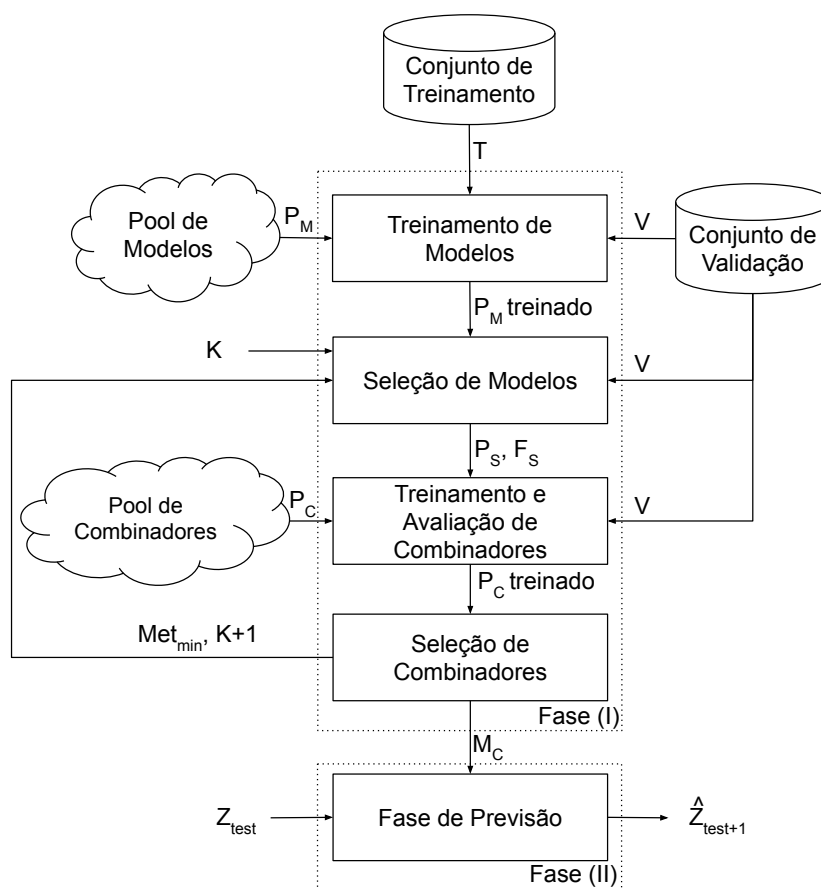
A seguir, Calatayud et al. (2022) propuseram a modelagem de ruído a partir de *stochastic differential equations*. Identificando a grande quantidade de ruídos em séries temporais relacionadas a crimes, se dispuseram a decompor uma série temporal relacionada a cidade de Valência na Espanha e observar em que medida o modelo replicaria com qualidade os dados reais, a partir da modelagem de ruídos [Calatayud et al. 2023].

No trabalho de Sinha et al. (2023) a abordagem para análise, resposta e prevenção do crime é projetada para identificar áreas propensas ao crime e descobrir pontos críticos com alta probabilidade de ocorrência. O trabalho proposto analisa ainda as tendências do crime ao longo do tempo para deduzir as horas mais seguras usando o objeto gráfico. Além disso, as tendências de ocorrência de crimes são previstas usando métodos simples de séries temporais e técnicas de suavização exponencial [Sinha and Biswas 2023].

Em todos os trabalhos observados nesta área de estudos, as pesquisas se dedicam a observar o desempenho de single models. Neste sentido, pretendemos oferecer alternativas para a melhoria dos modelos, utilizando a combinação a partir de ensembles estáticos ou por empilhamento (*stacking ensembles*), contribuindo, assim, com o avanço das pesquisas.

### 3. Método proposto

O método proposto é apresentado na Figura 1. A proposta pode ser dividida em duas fases: Fase de Treinamento (I) e Fase de Teste (II). Na Fase (I) a entrada do método é o pool  $P_M$  de modelos de previsão, o pool  $P_C$  de combinadores, a variável  $K$  e os conjuntos



**Figura 1. Fases de treinamento e teste do método proposto.**

de treinamento T e de validação V. A saída da Fase de Treinamento é a combinação de modelos selecionada. Essa combinação de modelos é empregada posteriormente na Fase de Teste. Nas próximas subseções as Fases de Treinamento e Teste serão descritas.

### 3.1. Fase de Treinamento

O objetivo na Fase de Treinamento é treinar o pool de modelos  $P_M$ , o pool de combinadores treináveis, avaliar e selecionar a melhor combinação de modelos para ser usada na Fase de Teste. Desse modo, o primeiro passo dessa fase é o treinamento do pool de modelos  $P_M$ . Para isso os conjuntos T e V são considerados para o treinamento dos modelos de Aprendizado de Máquina, enquanto que os conjuntos T e V são unidos para o treinamento dos modelos estatísticos. A saída da etapa de Treinamento dos modelos únicos é o pool  $P_M$  treinado. O pool  $P_M$  é avaliado na etapa de Seleção de Modelos com o objetivo de escolher o grupo dos K modelos  $P_S$  mais acurados. Essa avaliação é feita utilizando o conjunto V.

A previsão do conjunto de modelos selecionados  $F_S$  é utilizado como entrada para treinar o pool de combinadores  $P_C$ . Então, o grupo de combinadores treináveis de  $P_C$  são estimados utilizando a saída dos modelos selecionados  $F_S$  para gerar a melhor previsão para o conjunto V. O objetivo é gerar um desempenho melhor do que os combinadores não treináveis como média e mediana. A saída da etapa de Treinamento e Avaliação de Combinadores é o pool  $P_C$  treinado. É importante mencionar que em  $P_C$  também podem

haver combinadores que não necessitam de treinamento como média e mediana. O pool  $P_C$  é dado como entrada para o passo de Seleção de Combinadores. Nessa etapa o melhor sub-conjunto de preditores é selecionado como respectivo modelo de combinação. Essa seleção é realizada baseada em alguma métrica de desempenho. Após a etapa de seleção, o número de preditores é atualizado e uma nova etapa de treinamento de combinadores é iniciada. Caso o parâmetro  $K$  já tenha alcançado o valor máximo determinado pelo usuário, a fase de treinamento retorna a melhor combinação de modelos gerada  $M_C$  escolhida a partir do conjunto  $V$ .

### 3.2. Fase de Teste

Na Fase de Teste, que pode ser vista na Figura 1, o modelo  $M_C$  é utilizado para fazer a previsão do padrão de teste  $Z_{test}$ . Assim, para cada novo padrão,  $M_C$  gera a previsão para um passo à frente.

Nessa fase, a previsão é realizada em duas etapas. Primeiramente, os modelos de previsão que compõem  $M_C$  geram a primeira estimativa para  $Z_{test+1}$ . Após isso, o modelo de combinação recebe essas estimativas e gera a previsão final  $\hat{Z}_{test+1}$ .

## 4. Materiais e método

### 4.1. Banco de dados

Os experimentos foram conduzidos em três conjuntos de dados: crimes em Chicago, crime em Los Angeles e crimes violentos letais intencionais de Pernambuco (CVLI). Os dois primeiros conjuntos são frequentemente utilizados na literatura para a realização de experimentos [Safat et al. 2021, Tariq et al. 2021]. A série temporal referente aos crimes em Chicago reflete o relato de crimes (com exceção de assassinato) que ocorreram na cidade de Chicago de 2001 até o presente. Os dados são extraídos do sistema CLEAR (*Citizen Law Enforcement Analysis and Reporting*) do Departamento de Polícia de Chicago. Já a série temporal dos crimes de Los Angeles reflete incidentes de crimes de 2010 a 2019. São dados transcritos de relatórios de crimes originais, relatados em delegacias da respectiva cidade.

Por fim, a última série temporal se refere aos crimes violentos letais intencionais de Pernambuco (CVLI). CVLI é constituído pelos crimes de homicídio doloso, latrocínio, feminicídio, lesão corporal seguida de morte e outros crimes resultantes em mortes, excetuando-se os casos de morte por intervenção de agente do estado. Estes dados são coletados diariamente no IML, Polícia Científica, Coordenação de Plantão da Polícia Civil, Serviço de inteligência da PMPE e CIODS, sendo coletados e analisados pela Gerência Geral de Análise Criminal e Estatística (GGACE) de maneira preliminar, antes de que sejam disponibilizados de maneira pública. São dados que contemplam janeiro de 2004 em diante. Diferente das séries anteriores, esta série não relata os crimes contra patrimônio.

Embora seja possível realizar a coleta de dados até o ano de 2023, optou-se por utilizar os dados até o mês de dezembro de 2022. Esta decisão se dá, sobretudo, para evitar a possibilidade de dados faltantes, precavendo, portanto, possíveis falhas no modelo. Em todos os casos, os conjuntos são compostos por dados relacionados ao tipo de crime cometido, localização, características pessoais da vítima e data da ocorrência. Nem todas

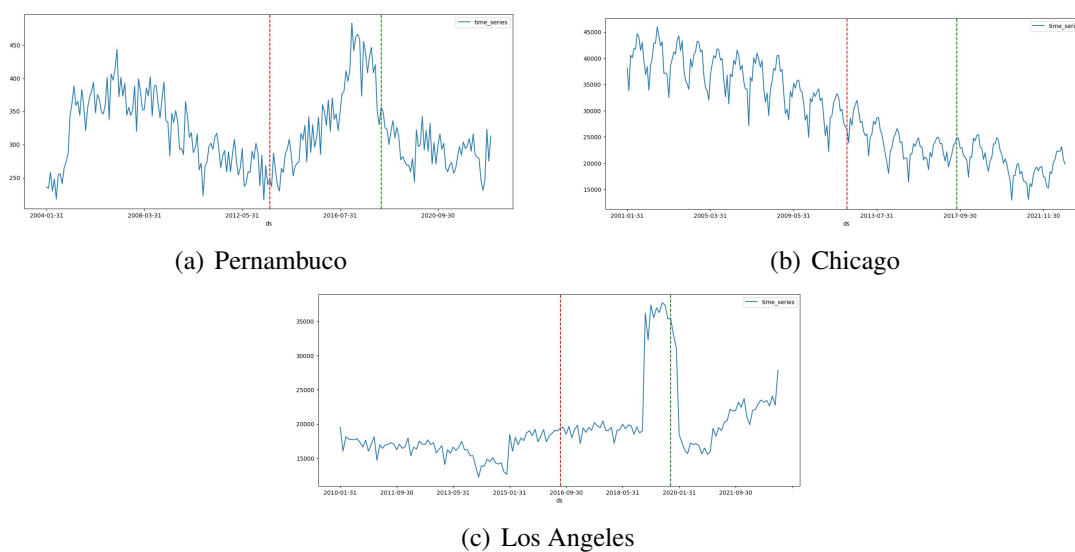
as informações são utilizadas, cabendo a nós, utilizar as informações de data e quantidade de crimes cometidos em um determinado período. Tabela 1 apresenta as três séries temporais.

**Tabela 1. Descrição da séries temporais de Pernambuco, Los Angeles e Chicago.**

Série temporal	Início	Fim	Tam.	Mín.	Máx.	Média	Mediana	DP
Pernambuco	01/2004	12/2022	228	217	483	318.4	310.0	55.4
Los Angeles	01/2010	12/2022	156	12220	37714	19392.4	18073.0	5242.8
Chicago	01/2001	12/2022	264	12918	46013	29212.4	28100.0	8344.5

## 4.2. Protocolo experimental

Para realização dos experimentos, as séries temporais foram divididos em conjuntos de treinamento, validação e teste, respectivamente com 50%, 25% e 25% das séries originais (Figura 2). Para o ARIMA, especificamente, como não é necessária a utilização do conjunto de validação, utilizou-se 75% para o treinamento e 25% para o teste. Os valores foram então normalizados para o intervalo  $[0.1; 0.9]$  com base nos valores máximos e mínimos dos conjuntos de treinamento e validação.



**Figura 2. Séries temporais e seus conjuntos de treinamento, validação e teste.**

Todos os modelos foram implementados em Python. Para treinamento dos modelos ARIMA, utilizou-se a função `auto.arima` do *package* `pmdarima` para busca dos melhores parâmetros. Por sua vez, para a definição dos parâmetros dos modelos de ML (SVR, MLP e LSTM), empregou-se o método *GridSearch* (Tabela 2). O treinamento do SVR foi realizado a partir do *package* `sklearn`. Por fim, os modelos MLP e LSTM foram treinados a partir do `tensorflow`. O algoritmo *Sliding Window* foi utilizado para janelamento das séries temporais, etapa necessária para treinamento dos modelos de ML. Foram realizadas previsões *one-step ahead* para todos os modelos.

A seguir foram implementados os modelos *ensembles*. Inicialmente foram criados os modelos estáticos de Médias Simples (do inglês *Simple Average* - SA) e Mediana

**Tabela 2. Parâmetros da GridSearch para os modelos SVR, MLP e LSTM.**

Modelos	Parâmetros	Valores
SVR	<i>Lag</i>	1, 3, 6, 9, 12, 15, 18
	Kernel	Linear, Radial, Sigmoid
	Gamma	1, 0.1, 0.01, 0.001
	Cost	0.1, 1, 10, 100, 1000, 10000
	Epsilon	0.1, 0.01, 0.001
MLP	<i>Lag</i>	1, 3, 6, 9, 12, 15, 18
	Função de ativação	RELU, Sigmoid
	Unidades na camada oculta	2, 5, 10, 15, 20
LSTM	Algoritmo	Adam
	<i>Lag</i>	1, 3, 6, 9, 12, 15, 18
	Função de ativação	RELU, Sigmoid
	Unidades na camada oculta	2, 5, 10, 15, 20
	Algoritmo	Adam

Simple (do inglês *Simple Median* - SM), os quais consideram a média e mediana das previsões dos modelos individuais. Foram também criados os *ensembles* eSVR e eMLP, os quais utilizam os modelos SVR e MLP, respectivamente, como meta-modelos que são treinados com base nas previsões dos modelos individuais. Por fim, foi criado o *ensemble* proposto.

### 4.3. Medidas de desempenho

Os modelos foram avaliados a partir de duas medidas de desempenho: *Root Mean Squared Error* (RMSE) e *Mean Absolute Percentage Error* (MAPE). Esta escolha se deu porque são avaliações com frequência utilizadas na área de previsão de séries temporais e especificamente na área de previsão de crimes. O RMSE é a raiz quadrada do erro médio. Por estar representado no intervalo da série temporal pode ser facilmente interpretada (Equação 1). Por sua vez, o MAPE representa o erro percentual absoluto médio (Equação 2). Diferentemente do RMSE, o MAPE é dimensionado em unidades percentuais.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (z_t - \hat{z}_t)^2}, \quad (1)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{z_t - \hat{z}_t}{z_t} \right|, \quad (2)$$

nas quais  $z_t$  é a observação no tempo  $t$ ,  $\hat{z}_t$  sua previsão e  $n$  o número de observações.

## 5. Resultados e discussão

A Tabela 3 apresenta as medidas de desempenho RMSE e MAPE para os modelos únicos e ensembles treináveis e não treináveis desenvolvidos considerados as três séries temporais de interesse. Para a abordagem que considera modelos únicos, os modelos SVR, ARIMA e MLP foram as melhores alternativas para as séries Pernambuco, Chicago e LA,

respectivamente. Nota-se que o modelo ARIMA foi a pior alternativa para as séries de Pernambuco e LA.

Já para a abordagem referente aos combinadores treináveis, os melhores modelos foram SA, SM<sub>3</sub> e SM. Neste caso, a performance das alternativas foi similar para as alternativas. Por sua vez, referente aos ensemble treináveis as melhores alternativas foram eSVR<sub>3</sub>, eMLP<sub>3</sub>, eSVR<sub>2</sub> e eSVR<sub>2</sub>. Os modelos que apresentaram melhor desempenho geral por série temporal foram eSVR<sub>3</sub> (Pernambuco), SM<sub>3</sub> (Chicago), MLP e SM (Los Angeles).

A Tabela 4 apresenta o rank dos modelos implementados. Neste caso, os três melhores, de acordo com a média e mediana, foram SA, SM e SM<sub>3</sub>. De forma geral, o ranking médio dos ensembles não treináveis foi melhor do que para os os ensembles treináveis. Já considerando as diferentes abordagens, os modelos únicos foram os piores.

**Tabela 3. Medidas de desempenho para os modelos individuais e ensembles não treináveis e treináveis. Valores em negrito indicam o melhor valor para cada tipo de abordagem. Os valores em negrito e sublinhado indicam o melhor desempenho entre os modelos avaliados.**

Abordagem	Modelos	Pernambuco		Chicago		LA	
		RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
Modelos Únicos	ARIMA	32.69	6.87	<b>2173.80</b>	<b>7.28</b>	5003.23	8.93
	SVR	<b>24.11</b>	<b>6.78</b>	2975.28	13.55	2556.44	6.33
	MLP	24.24	6.79	3011.97	13.61	<b><u>1376.45</u></b>	5.00
	LSTM	24.73	6.94	2305.90	9.65	1445.02	5.06
Ensemble Não Treinável	SA	<b>23.99</b>	<b>6.63</b>	1795.13	7.41	1661.60	5.00
	SM	24.20	6.77	1801.77	7.38	<b>1467.66</b>	<b>4.61</b>
	SA <sub>2</sub>	24.17	6.79	1994.58	8.32	1971.03	5.56
	SA <sub>3</sub>	24.35	6.84	1748.88	7.32	1749.21	5.11
	SM <sub>2</sub>	24.17	6.79	1994.58	8.32	1971.03	5.56
	SM <sub>3</sub>	24.39	6.86	<b><u>1614.38</u></b>	<b>6.91</b>	1492.21	4.65
Ensemble Treinável	eSVR	23.59	6.44	2295.99	10.11	2145.35	<b>7.32</b>
	eMLP	24.22	6.86	2037.20	8.75	2579.05	9.20
	eSVR <sub>2</sub>	23.79	6.58	2347.02	10.24	<b>2117.74</b>	7.39
	eSVR <sub>3</sub>	<b><u>23.30</u></b>	<b>6.36</b>	2345.93	10.30	2442.98	8.28
	eMLP <sub>2</sub>	23.95	6.78	2007.17	8.54	2773.91	9.96
	eMLP <sub>3</sub>	24.20	6.85	<b>1782.30</b>	<b>7.45</b>	2599.08	9.27
Proposto		23.79	6.58	2345.93	10.30	2117.74	7.39

Em relação ao método proposto, o seu desempenho foi bom para a série de Pernambuco (ranking médio 3), mas foi ruim para Chicago e LA. Especificamente para a série de Pernambuco, o RMSE da validação indica (Tabela 5) que a melhor alternativa de ensemble é aquele que combina os dois melhores modelos únicos por meio de um SVR (eSVR<sub>2</sub>). Entretanto, nota-se que o RMSE dessa alternativa (eSVR<sub>2</sub>) e o da que foi melhor no teste (eSVR<sub>3</sub>) foram muito próximos, sendo 0.1102 e 0.1108, respectivamente. De forma similar, para a série de Chicago, a seleção com base na validação indica que o melhor modelo é o eSVR<sub>3</sub>, quando foi o SM<sub>3</sub> que obteve melhor desempenho no conjunto de teste (Tabela 3). Resultado similar ocorre para série de LA.



**Tabela 4. Rank dos modelos considerando as medidas de desempenho RMSE e MAPE. Colunas à direita apresentam a estatística do rank por modelo.**

Modelos	Pernambuco		Chicago		LA		Estatísticas		
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	Média	Mediana	DP
ARIMA	17	16	10	2	17	14	12.7	15.0	5.9
SVR	7	8	16	16	13	9	11.5	11.0	4.0
MLP	13	11	17	17	<b>1</b>	3	10.3	12.0	6.9
LSTM	16	17	12	11	2	5	10.5	11.5	6.0
SA	6	5	4	5	5	4	4.8	5.0	0.8
SM	10	6	5	4	3	<b>1</b>	4.8	4.5	3.1
SA <sub>2</sub>	8	9	6	7	7	7	7.3	7.0	1.0
SA <sub>3</sub>	14	12	2	3	6	6	7.2	6.0	4.8
SM <sub>2</sub>	8	9	6	7	7	7	7.3	7.0	1.0
SM <sub>3</sub>	15	14	<b>1</b>	<b>1</b>	4	2	6.2	3.0	6.6
eSVR	2	2	11	12	11	10	8.0	10.5	4.7
eMLP	12	15	9	10	14	15	12.5	13.0	2.6
eSVR <sub>2</sub>	3	3	15	13	9	11	9.0	10.0	5.1
eSVR <sub>3</sub>	<b>1</b>	<b>1</b>	13	14	12	13	9.0	12.5	6.2
eMLP <sub>2</sub>	5	7	8	9	16	17	10.3	8.5	5.0
eMLP <sub>3</sub>	11	13	3	6	15	16	10.7	12.0	5.2
Proposto	3	3	13	14	9	11	8.8	10.0	4.8

Logo, para os casos em questão, a seleção da estratégia de combinação com base no RMSE da validação tende a ser problemática quando os conjuntos de validação não forem representativos o suficiente, ou por possuírem poucas observações, ou por possuírem características diferentes do conjunto de teste.

**Tabela 5. RMSE do conjunto de validação utilizado para seleção do K modelos. O RMSE foi calculado com base nos valores normalizados.**

Abordagem	Modelo	PE	CH	LA
Ensemble Não Treinável	SA <sub>2</sub>	0.1137	0.0344	0.1594
	SA <sub>3</sub>	0.1128	0.0330	0.1586
	SM <sub>2</sub>	0.1137	0.0344	0.1594
	SM <sub>3</sub>	0.1134	0.0344	0.1614
Ensemble Treinável	eSVR <sub>2</sub>	<b>0.1102</b>	0.0306	<b>0.1539</b>
	eSVR <sub>3</sub>	0.1108	<b>0.0305</b>	0.1662
	eMLP <sub>2</sub>	0.1157	0.0330	0.3693
	eMLP <sub>3</sub>	0.1167	0.0371	0.3434

Embora o modelo proposto, no geral, apresente um resultado abaixo dos outros ensembles abordados no experimento completo, esta pesquisa avança no que concerne a literatura de previsão de crimes. A utilização de modelos únicos, como realizado em trabalhos anteriores, [Safat et al. 2021, Tariq et al. 2021], apresentam uma limitação frente a modelos combinados.

Partindo da premissa que a adoção de um ensemble é promissora porque reduz o risco de selecionar modelos super ou subajustados, equilibra o *trade-off* entre viés e variância e combina os pontos fortes de diferentes modelos [Zhang 2007], como apontado

em sessões anteriores, mostramos que há muito a se avançar na discussão de previsão de crimes. Modelos únicos podem ser úteis, mas, no geral, tendem a ser superados por modelos combinados, inclusive, quando tratamos de *crime forecasting*.

## 6. Conclusão

A previsão de crimes é uma tarefa desafiadora devido à complexidade das séries temporais envolvidas. Logo, para lidar com essa complexidade e capturar a natureza dinâmica dos padrões temporais em constante mudança, é necessário o uso de métodos de previsão adequados. O presente artigo apresentou diferentes métodos de combinação comumente utilizados na literatura, como SA, SM, eSVR e eMLP, e propôs uma nova abordagem ensemble a fim de aumentar a acurácia das previsões de séries temporais de crimes.

Com base nos resultados obtidos, observou-se que os modelos ensembles demonstraram um desempenho preditivo superior em comparação aos modelos únicos utilizados, como pressuposto a *a priori*. No entanto, em relação à abordagem proposta, conclui-se que a seleção de um subconjunto de modelos únicos com base na validação pode apresentar desafios quando os conjuntos de validação não são suficientemente representativos. Isso pode ocorrer devido ao número limitado de observações disponíveis ou às características distintas em relação ao conjunto de teste. Portanto, é importante considerar essas limitações ao implementar a estratégia de seleção de modelos para garantir a robustez e a confiabilidade das previsões.

Duas limitações podem ser discutidas na execução deste trabalho: o número reduzido de séries temporais de crimes (apenas três) e o tamanho limitado dessas séries. Portanto, como perspectiva para trabalhos futuros, sugere-se a utilização de séries temporais adicionais, provenientes de diferentes localidades e com características diversas. Além disso, em relação ao método proposto, pretende-se desenvolver uma nova estratégia para a seleção dos modelos individuais a serem combinados, visando aprimorar o potencial preditivo do mesmo.

## Referências

- Anuvarshini, S., Nidhi, D., Sree, D., et al. (2022). Crime forecasting: A theoretical approach. In *2022 IEEE 7th International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, volume 7, pages 37–41. IEEE.
- Calatayud, J., Jornet, M., and Mateu, J. (2023). Modeling noisy time-series data of crime with stochastic differential equations. *Stochastic Environmental Research and Risk Assessment*, 37(3):1053–1066.
- Chen, P., Yuan, H., and Shu, X. (2008). Forecasting crime using the arima model. In *2008 fifth international conference on fuzzy systems and knowledge discovery*, volume 5, pages 627–630. IEEE.
- Cohen, L. E. (1981). Modeling crime trends: A criminal opportunity perspective. *Journal of Research in Crime and Delinquency*, 18(1):138–164.
- Detotto, C. and Otranto, E. (2010). Does crime affect economic growth? *Kyklos*, 63(3):330–345.
- Kelly, M. (2000). Inequality and crime. *Review of economics and Statistics*, 82(4):530–539.

- Khairuddin, A. R., Alwee, R., and Haron, H. (2019). A review on applied statistical and artificial intelligence techniques in crime forecasting. In *IOP conference series: materials science and engineering*, volume 551, page 012030. IOP Publishing.
- McDowall, D., Loftin, C., and Pate, M. (2012). Seasonal cycles in crime, and their variability. *Journal of quantitative criminology*, 28:389–410.
- Olligschlaeger, A. M. (1997). Artificial neural networks and crime mapping. *Crime mapping and crime prevention*, 1:313.
- Safat, W., Asghar, S., and Gillani, S. A. (2021). Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques. *IEEE access*, 9:70080–70094.
- SILVA, E. G. d. et al. (2021). Uma abordagem de seleção dinâmica de preditores baseada nas janelas temporais mais recentes.
- Sinha, B. B. and Biswas, T. (2023). An efficient framework for forecasting of crime trend using machine learning technique. In *Proceedings of International Conference on Data Science and Applications: ICDSA 2022, Volume 2*, pages 741–755. Springer.
- Tariq, H., Hanif, M. K., Sarwar, M. U., Bari, S., Sarfraz, M. S., and Oskouei, R. J. (2021). Employing deep learning and time series analysis to tackle the accuracy and robustness of the forecasting problem. *Security and Communication Networks*, 2021:1–10.
- Veiga, J. E. D. (2014). O âmago da sustentabilidade. *estudos avançados*, 28:7–23.
- Zhang, G. P. (2007). A neural network ensemble method with jittered training data for time series forecasting. *Information Sciences*, 177(23):5329–5346.
- Zhuang, Y., Almeida, M., Morabito, M., and Ding, W. (2017). Crime hot spot forecasting: A recurrent model with spatial and temporal information. In *2017 IEEE International Conference on Big Knowledge (ICBK)*, pages 143–150. IEEE.