

Enhancing Crime Forecasting with Hybrid Approaches: Integrating Machine Learning and Statistical Models

Paulo Cauas¹, Ruam Pastor¹, Paulo S. G. de Mattos Neto¹, Filipe C. de L. Duarte²

¹Centro de Informática – Universidade Federal de Pernambuco (CIN/UFPE)
Caixa Postal 50740-560 - Av. Jorn. Aníbal Fernandes, s/n - Cidade Universitária,
Recife - PE, Brazil

²Departamento de Finanças e Contabilidade – Universidade Federal da Paraíba (UFPB)
João Pessoa-PB, Brasil.

Abstract. *Crime forecasting represents a significant challenge for public security. This article investigates the use of hybrid systems for crime time series forecasting, utilizing data from different Brazilian states. The proposed approach consists of combining Machine Learning (ML) models to capture non-linear patterns in the time series, followed by the application of the statistical Autoregressive Integrated Moving Average (ARIMA) model to the forecast residuals, aiming to model remaining linear patterns. For evaluation, the proposed hybrid versions employed the following ML models: Long Short-Term Memory (LSTM), Multilayer Perceptron (MLP), and Support Vector Regression (SVR). The proposed versions (LSTM+ARIMA, MLP+ARIMA, and SVR+ARIMA) were compared with individual models and traditional hybrid approaches from the literature. The results demonstrate that the proposed hybrid approaches achieve superior performance, especially in scenarios with high complexity and a predominance of nonlinear patterns, reinforcing the importance of strategies that integrate different modeling paradigms. This study advances the field of crime forecasting by demonstrating the effectiveness of the proposed hybrid approach. The findings highlight the potential of hybrid systems as a robust and innovative foundation for decision support systems in public security, showing that the integration of different modeling paradigms can offer substantial gains in both accuracy and practical utility.*

Resumo. *A previsão de crimes representa um desafio significativo para a segurança pública. Este artigo investiga o uso de sistemas híbridos para previsão de séries temporais de crimes, utilizando dados de diferentes estados brasileiros. A proposta consiste em combinar modelos de Aprendizado de Máquina (AM) para capturar padrões não lineares das séries temporais, seguidos pela aplicação do modelo estatístico Autoregressive Integrated Moving Average (ARIMA) sobre os resíduos de previsão, visando modelar padrões lineares remanescentes. Para avaliação, as versões híbridas propostas utilizaram os modelos de AM: Long Short-Term Memory (LSTM), Multilayer Perceptron (MLP) e Support Vector Regression (SVR). As versões da abordagem proposta (LSTM+ARIMA, MLP+ARIMA e SVR+ARIMA) foram comparadas com modelos individuais e híbridos tradicionais da literatura. Os resultados demonstram que as abordagens híbridas propostas apresentam desempenho superior, especialmente em cenários com alta complexidade e predominância de padrões*

não lineares, reforçando a importância de estratégias que integrem diferentes paradigmas de modelagem. O estudo avança o campo da previsão de crimes ao demonstrar a eficácia da abordagem híbrida proposta. Os resultados reforçam o potencial dos sistemas híbridos como uma base robusta e inovadora para sistemas de apoio à decisão em segurança pública, demonstrando que a integração de diferentes paradigmas de modelagem pode oferecer ganhos substanciais em precisão e utilidade prática.

1. Introdução

A previsão de séries temporais desempenha um papel crucial em diversas áreas, desde aplicações sociais [Duarte et al. 2024], econômicas [Khashei and Hajirahimi 2019], segurança pública [Safat et al. 2021], engenharia até a gestão de recursos naturais e de energia [Izidio et al. 2021, Shi et al. 2024], fornecendo subsídios para a tomada de decisão e a formulação de políticas públicas eficazes.

A previsão de crimes visa analisar dados históricos para identificar padrões e tendências, gerando previsões que possam auxiliar na antecipação de futuras ocorrências delitivas [Mandalapu et al. 2023]. Contudo, ela configura-se como um problema particularmente desafiador devido à complexa interação de múltiplos fatores, incluindo dinâmicas sociais, espaciais e temporais que influenciam a ocorrência de delitos [Dakalbab et al. 2022]. As séries temporais de crimes frequentemente exibem características como não estacionariedade, flutuações aleatórias, mudança de conceito e sazonalidade, dificultando sua modelagem por abordagens tradicionais [Chainey and Tompson 2008, Safat et al. 2021, Mandalapu et al. 2023]. Nesse contexto, embora existam limitações dos métodos tradicionais de previsão em capturar padrões complexos não lineares, modelos lineares foram aplicados com o objetivo de, ao menos compreender o comportamento dos dados [İlgün, Esen Gül and Dener 2025]. Esses modelos têm como grande vantagem uma forma bem consolidada para estimação de seus parâmetros, diagnóstico de resíduos e avaliação de desempenho, o que contribui para maior robustez e confiabilidade das previsões [de Oliveira et al. 2021]. Entre os modelos estatísticos, o que se destaca em utilização é o modelo *Autoregressive Integrated Moving Average* (ARIMA) e sua derivação sazonal (*Seasonal ARIMA* - SARIMA), recorrentemente utilizados nesta tarefa [Muthamizharasan and Ponnusamy 2024, Noor et al. 2022, Safat et al. 2021].

Nos últimos anos, trabalhos da literatura têm aplicado técnicas de Aprendizado de Máquina (AM) [Dakalbab et al. 2022, Shah et al. 2021]. Modelos de AM, como *Support Vector Regression* (SVR) [Alwee et al. 2013, Khairuddin et al. 2020], *Random Forests* [Abdulraheem et al. 2022] e RNAs clássicas, como *Multilayer Perceptron* (MLP), e de aprendizado profundo (*Deep Learning* - DL) como *Long Short-Term Memory* (LSTM) [Mao et al. 2025], demonstraram uma capacidade notável de extrair padrões complexos, não lineares e interações sutis a partir de vastos conjuntos de dados criminais [Kang and Kang 2017, Safat et al. 2021, Mandalapu et al. 2023, İlgün, Esen Gül and Dener 2025]. Os pontos fortes da AM residem em sua abordagem orientada a dados, que permite a descoberta de relações sem a necessidade de uma especificação formal do modelo e sua flexibilidade em adaptar-se a diferentes tipos de crimes e contextos geográficos [Dakalbab et al. 2022, İlgün, Esen Gül and Dener 2025]. Contudo, os modelos de AM, em sua maioria, também apresentam fraquezas,

como a necessidade de volumes de dados de alta qualidade para treinamento eficaz [Mandalapu et al. 2023], o risco de sobreajuste (*overfitting*), subajuste (*underfitting*) e má especificação aos dados de treinamento [de Oliveira et al. 2021].

Na literatura, sistemas híbridos, que utilizam a modelagem de resíduos¹, têm sido propostos com o objetivo de superar as deficiências e limitações inerentes dos modelos monolíticos, sejam fundamentados em abordagens estatísticas tradicionais ou em AM [Zhang 2003, Khashei and Hajirahimi 2019, Duarte et al. 2024]. Esses sistemas têm se consolidado como uma solução robusta e eficaz, pois possibilitam a combinação das vantagens dos modelos estatísticos e de AM, permitindo que os pontos fortes de cada componente mitigue as limitações da outra [Zhang 2003, Khashei and Hajirahimi 2019, Duarte et al. 2024]. Um sistema híbrido é composto por três etapas sequenciais: (i) modelagem da série temporal utilizando um modelo de previsão estatístico linear ou de AM; (ii) modelagem da série de resíduos empregando um modelo de natureza distinta da etapa (i); e (iii) combinação das previsões obtidas em (i) e (ii) por meio de alguma função de combinação, sendo a soma a mais comumente utilizada [Pai and Lin 2005, de Oliveira et al. 2021].

A maioria dos trabalhos na literatura [Dave et al. 2021, Manowska et al. 2021, Liu et al. 2021, Zhang et al. 2024] emprega modelos estatísticos para a modelagem da série temporal e modelos de AM para a modelagem dos resíduos, respectivamente. Dessa forma, essa combinação foca na modelagem dos padrões lineares da série temporal por meio dos modelos estatísticos. No entanto, a modelagem dos resíduos apresenta desafios, pois as séries de erro são heterocedásticas e podem conter não apenas estruturas não lineares, mas também ruído aleatório [de Oliveira et al. 2021]. Essa característica dificulta a identificação de relações temporais relevantes, aumenta o risco de *overfitting* e pode comprometer a acurácia dos modelos de AM, consequentemente afetando o desempenho de todo o sistema híbrido. Alternativamente, alguns trabalhos na literatura [Khashei and Hajirahimi 2019, Hajirahimi and Khashei 2022] sugerem que inverter a ordem de modelagem - ou seja, aplicar primeiro o modelo de aprendizado de máquina e, em seguida, um modelo estatístico linear sobre os resíduos - pode gerar resultados mais frutíferos em termos de acurácia.

Este artigo propõe uma nova abordagem para a previsão de séries temporais de crimes, combinando modelos de aprendizado de máquina com modelos estatísticos lineares por meio da modelagem da série de resíduos. A abordagem proposta tem como objetivo:

- Abordar de forma eficiente o desafio da seleção de modelos na tarefa de previsão de crimes;
- Maximizar a acurácia dos modelos de AM, focando na modelagem de padrões complexos da série temporal;
- Permitir que o modelo estatístico, ajustado sobre os resíduos, complemente o modelo de AM ao capturar padrões lineares e autocorrelações residuais, reduzindo o erro de previsão;
- Combinar a habilidade dos modelos de AM em capturar padrões complexos com a força dos modelos estatísticos em modelar padrões lineares, alcançando melhor

¹Resíduos, ou série de erros, são definidos como as diferenças entre os valores observados da série temporal e os valores previstos por um modelo.

acurácia que os modelos utilizados separadamente.

A abordagem proposta é composta por três fases: 1) modelagem da série temporal utilizando um modelo de AM; 2) modelagem dos resíduos; 3) combinação das previsões obtidas nas fases (1) e (2). A abordagem emprega na fase (1) três modelos de AM bem conhecidos na literatura: LSTM, MLP e SVR. Na fase (2), o tradicional modelo ARIMA, juntamente com a metodologia de Box & Jenkins, é utilizado na previsão dos resíduos.

O artigo está estruturado em cinco seções, além desta Introdução. Na Seção 2, é apresentada uma revisão da literatura sobre previsão de crimes e modelagem híbrida. A Seção 3 descreve a abordagem proposta. A Seção 4 descreve o protocolo experimental adotado para avaliação dos modelos, detalhando os dados utilizados e as métricas de avaliação. Na Seção 5, os resultados são apresentados e analisados. Por fim, a Seção 6 conclui o artigo, abordando suas limitações e sugerindo direções para pesquisas futuras.

2. Revisão da literatura

Desde os anos 1990, a previsão de crimes tornou-se estratégica para a segurança pública, impulsionada por iniciativas como as do *National Institute of Justice* (NIJ) [Olligschlaeger 1997, Gorr and Harries 2003]. Apesar do avanço, o desafio persiste devido à natureza multifacetada do crime, influenciado por fatores socioeconômicos, urbanos e ambientais [Resende and Andrade 2011, Xu et al. 2018, Lab and Hirschel 1988]. Os padrões criminais variam no tempo e espaço, apresentando componentes não lineares e sazonalidades [İlgün, Esen Gül and Dener 2025, McDowall et al. 2012], o que exige métodos adaptativos e flexíveis — cenário ideal para abordagens híbridas e aprendizado de máquina.

A previsão univariada de séries temporais de crime surgiu como uma alternativa prática para lidar com parte dessas dificuldades [Gorr and Harries 2003]. A modelagem preditiva baseada exclusivamente em dados históricos do fenômeno em análise, como as taxas de criminalidade, contorna a dependência de informações exógenas frequentemente de difícil aquisição. Essa abordagem permite a identificação de padrões recorrentes – a exemplo de elevações sazonais nos índices de crimes durante períodos festivos – viabilizando, assim, a geração de previsões mais precisas.

Embora a previsão de séries temporais de crimes tenha sido explorada com modelos estatísticos e de AM a complexidade dos dados criminais, que frequentemente conjugam padrões lineares e não lineares, representa um desafio. [Alwee et al. 2013] propuseram um modelo híbrido que integrou um SVR e um ARIMA, otimizados a partir do algoritmo *Particle Swarm Optimization* (PSO), para prever taxas anuais de crimes contra propriedade nos Estados Unidos. Comparando a proposta com os modelos individuais, utilizando ou não o PSO, eles concluíram que a hibridização resultou em maior acurácia de previsão para diversas métricas. [Marzan et al. 2017] compararam os modelos Regressão Linear (RL), Regressão de Processos Gaussianos (RPG), *Sequential Modeling Regression* (SMOReg) e *Multilayer Perceptron* (MLP) por meio da métrica *Mean Absolute Percentage Error* (MAPE) para dados diários e semanais de crimes de dezesseis distritos da capital da Filipinas. Para a frequência diária, os resultados foram mistos, enquanto que para os dados semanais, o MLP demonstrou superioridade. [Safat et al. 2021] aplicaram os modelos *Long Short-Term Memory* (LSTM) e ARIMA para classificação e previsão de séries temporais, respectivamente. Entretanto, não ap-

resentaram comparação com outros modelos de previsão. [Pastor et al. 2023] comparou modelos individuais (ARIMA, MLP, SVR e LSTM), combinações pela média e mediana, e combinações treináveis utilizando MLP e SVR como meta modelos para a previsão de taxas mensais de crimes de Chicago, Los Angeles e Pernambuco. Os resultados foram mistos, pois não houve um modelo que se sobressaiu ao longo dos experimentos. [İlgün, Esen Gül and Dener 2025] avaliou o desempenho dos modelos Holt-Winters, Prophet, ARIMA, LSTM e *Bidirectional*-LSTM (BiLSTM) na tarefa de previsão de taxas de crimes de diversos distritos da Chicago, Filadélfia e São Francisco. Os resultados mostraram que a rede neural BiLSTM foi superior para a maioria dos distritos policiais.

Diante disso, notou-se que os modelos individuais dificilmente alcançam desempenho ótimo universalmente, evidenciando limitações na captura simultânea dessas distintas características. Assim, uma vez que os dados são complexos e os modelos individuais raramente se sobressaem em todos os cenários, as combinações de modelos tornam-se mais interessantes, pois permitem explorar diversas características dos dados de forma concorrente, visando um desempenho preditivo superior [Alwee et al. 2013, Pastor et al. 2023].

3. Abordagem Proposta

A Figura 1 ilustra a abordagem híbrida proposta. A abordagem é composta por duas fases: (I) Treinamento e (II) Teste. A fase (I) recebe como entrada o conjunto de treinamento da série temporal Z_{tr} que será utilizado para treinar os modelos de AM (M_{AM}) e estatístico (M_{ES}). A saída da fase de Treinamento são os modelos treinados: (M_{AM}) e (M_{ES}). Na fase (II), as entradas consistem nos padrões de teste da série temporal z_q e dos resíduos e_q . A saída da fase de Teste é a previsão $\hat{z}_{q+h}$, onde h representa o horizonte de previsão escolhido. As seções a seguir apresentam uma descrição mais detalhada das fases de Treinamento e Teste.

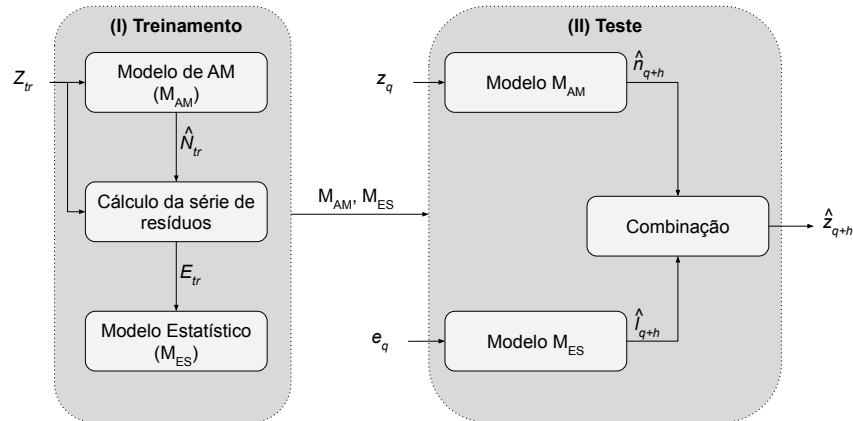


Figure 1. Fases da abordagem proposta: (I) Treinamento e (II) Teste.

3.1. Fase de Treinamento

Durante a fase de Treinamento, a abordagem proposta é composta por três passos: (i) treinamento do M_{AM} utilizando Z_{tr} e geração da previsão $\hat{N}_{tr}$; (ii) estimação da série de resíduos (E_{tr}); e (iii) modelagem de E_{tr} utilizando o modelo linear M_{ES} . A série de erros (E_{tr}) é calculada pela diferença entre Z_{tr} e $\hat{N}_{tr}$.

3.2. Fase de Teste

Na fase de Teste, a abordagem proposta recebe os padrões de teste da série (z_q) e dos resíduos (e_q) como entrada. Em seguida, os modelos M_{AM} e M_{ES} geram as previsões para a série temporal $\hat{n}_{q+h}$ e os resíduos $\hat{l}_{q+h}$, respectivamente. Por fim, a previsão final $\hat{z}_{q+h}$ é calculada a partir da soma das saídas das duas etapas anteriores, assim $\hat{z}_{q+h} = \hat{n}_{q+h} + \hat{l}_{q+h}$.

4. Protocolo Experimental

Este estudo utilizou séries temporais mensais de dois tipos de crimes: Mortes Violentas Intencionais (MVI) e Crimes contra o Patrimônio (CVP). Para MVI, foram analisadas séries temporais de três estados brasileiros: Pernambuco, Alagoas e Ceará. Para CVP, foram utilizadas séries de Pernambuco e Ceará. Todas as séries temporais foram divididas em três conjuntos: treinamento (50%), validação (25%) e teste (25%). O conjunto de treinamento foi utilizado para ajustar os parâmetros dos modelos. O conjunto de validação serviu evitar o *overfitting* dos modelos de AM e para a seleção dos melhores hiperparâmetros. O conjunto de teste foi reservado exclusivamente para a avaliação final dos modelos.

Todas as séries temporais foram normalizadas utilizando o método *Min-Max Scaler*, que transforma os valores para o intervalo $[0,1]$. Esta normalização foi aplicada para facilitar o treinamento dos modelos de AM, especialmente as redes neurais, que são sensíveis à escala dos dados.

Três versões da abordagem proposta foram desenvolvidas variando os modelos de AM utilizados na modelagem da série temporal: LSTM, MLP e SVR. O ARIMA foi empregado na modelagem dos resíduos de todas as três versões. As versões propostas são denotadas como LSTM+ARIMA, MLP+ARIMA e SVR+ARIMA. Para comparação foram implementados 4 modelos individuais e 3 sistemas híbridos. Os modelos individuais foram: ARIMA [İlgün, Esen Gül and Dener 2025], LSTM [Safat et al. 2021], MLP [Marzan et al. 2017] e SVR [Alwee et al. 2013]. Sistemas híbridos da literatura que utilizam a modelagem de resíduos também foram avaliados, sendo eles ARIMA+MLP [Zhang 2003, Khashei and Hajirahimi 2019], ARIMA+LSTM [Dave et al. 2021, Zhang et al. 2024] e ARIMA+SVR [Zhang et al. 2019, Rubio and Alba 2022].

Para otimização dos hiperparâmetros, foi utilizado o método de busca em grade para os modelos MLP, SVR e LSTM, avaliando os resultados por meio do Erro Quadrático Médio (*Mean Squared Error* - MSE). A Tabela 1 apresenta os hiperparâmetros avaliados para cada modelo.

Para avaliar o desempenho preditivo dos modelos, foram utilizadas duas métricas de erro no conjunto de teste: Raiz do Erro Quadrático Médio (*Root Mean Squared Error* - RMSE) e Erro Médio Absoluto (*Mean Absolute Error* - MAE). O RMSE (Equação 1) quantifica a magnitude média dos erros, penalizando mais fortemente os erros maiores devido ao termo quadrático. O RMSE foi escolhido por sua maior sensibilidade a erros grandes [Hodson 2022], o que é particularmente relevante em previsões de crimes, onde subestimar significativamente um pico de criminalidade pode ter consequências graves

Table 1. Parâmetros utilizados na busca em grade para cada modelo.

Modelo	Parâmetro	Valores Testados
MLP	Quantidade de neurônios por camada escondida	(100,), (50, 50), (30, 30, 30)
	Função de ativação	relu, tanh, logistic
	Taxa de aprendizagem	0.001, 0.01, 0.1
	Otimizador	adam, sgd, lbfgs
LSTM	Neurônios - camada escondida	5, 10, 50, 100, 500
	Função de ativação	relu, sigmoid
	Otimizador	adam
SVR	Kernel	linear, RBF, sigmoid
	Função de custo	0,1, 1, 10
	Epsilon	0,1, 0,01, 0,001
	Gamma	1, 0,1, 0,01, 0,001

para políticas públicas.

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

O MAE (Equação 2) mede a magnitude média dos erros absolutos da variável de interesse, sem considerar sua direção, facilitando sua interpretabilidade [Wang and Lu 2018].

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |z_t - \hat{z}_t|. \quad (2)$$

Nas Equações 1 e 2 o valor real é representado por z_t , $\hat{z}_t$ é o valor previsto e n é o número de observações.

A média da diferença percentual do RMSE entre os modelos par a par foi calculada para avaliar os resultados relativos dos modelos. A diferença percentual do RMSE do modelo i em relação ao modelo j para uma série temporal Z é dada por:

$$\Delta_{i,j}^{(Z)} = \frac{\text{RMSE}_j^{(Z)} - \text{RMSE}_i^{(Z)}}{\text{RMSE}_j^{(Z)}} \times 100. \quad (3)$$

Assim, quanto maior o valor de $\Delta_{i,j}^{(Z)}$, melhor é o modelo i em relação ao modelo j . Nesse trabalho, a média da diferença percentual do modelo i em relação ao modelo j é calculada considerando todas as séries temporais analisadas, visando gerar um comparativo geral entre os modelos.

5. Resultados

A Tabela 2 apresenta uma comparação entre as versões da abordagem proposta, os modelos monolíticos e os sistemas híbridos da literatura na previsão das séries temporais de MVI para os estados de Pernambuco, Alagoas e Ceará. Em todos os conjuntos de dados avaliados, pelo menos uma das três versões da abordagem proposta obteve o melhor desempenho em termos de RMSE e MAE. Especificamente, o modelo LSTM+ARIMA destacou-se em Pernambuco, enquanto o MLP+ARIMA apresentou os melhores resultados para o Ceará. Para Alagoas, as combinações MLP+ARIMA e SVR+ARIMA atingiram, respectivamente, os menores valores de RMSE e MAE. De modo geral, a modelagem

residual contribuiu para aprimorar as previsões do modelo de AM. No entanto, em 3 de 9 casos analisados, o modelo linear não conseguiu melhorar a previsão da série temporal, evidenciando o desafio inerente à modelagem dos resíduos. Esses resultados ressaltam a importância de investigar a tarefa de seleção de modelos para essa tarefa, visando melhorar a modelagem dos resíduos.

Table 2. Resultados dos modelos avaliados em termos de RMSE e MAE para as séries temporais de Mortes Violentas Intencionais (MVI). O melhor resultado está destacado em negrito.

Abordagem	Modelo	Pernambuco		Alagoas		Ceará	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
Modelos Individuais	ARIMA	0,139	0,120	0,197	0,145	0,137	0,118
	LSTM	0,081	0,065	0,173	0,140	0,078	0,061
	MLP	0,080	0,067	0,131	0,104	0,077	0,062
	SVR	0,080	0,067	0,127	0,093	0,138	0,118
Sistemas Híbridos da literatura	ARIMA+LSTM	0,134	0,115	0,203	0,148	0,141	0,121
	ARIMA+MLP	0,142	0,123	0,212	0,157	0,127	0,110
	ARIMA+SVR	0,132	0,113	0,248	0,177	0,137	0,117
Proposta	LSTM+ARIMA	0,070	0,057	0,162	0,128	0,083	0,066
	MLP+ARIMA	0,082	0,068	0,121	0,089	0,074	0,060
	SVR+ARIMA	0,076	0,063	0,122	0,086	0,272	0,253

A Tabela 3 mostra a comparação entre as versões da abordagem proposta, os modelos monolíticos e os sistemas híbridos da literatura na previsão das séries temporais de CVP de Pernambuco e Ceará. Mais uma vez, a abordagem proposta alcançou uma acurácia de destaque quando comparada com os demais modelos. Todas as abordagens propostas melhoraram o modelo de AM inicial nos dois casos, com destaque para a série do Ceará, na qual a melhoria alcançou uma ordem de grandeza. A combinação LSTM+ARIMA obteve o melhor desempenho nas duas bases de dados, melhorando o RMSE do LSTM em uma ordem de grandeza em ambos os casos. Para as séries de CVP, o ARIMA obteve melhor acurácia que os modelos de AM individuais, contrastando com o cenário da série de MVI. Apesar disso, a combinação ARIMA+MLP foi capaz de melhorar a acurácia do modelo inicial. Nos demais casos, os modelos de AM tiveram dificuldade em modelar os resíduos do conjunto de treinamento provenientes do ARIMA.

Table 3. Resultados dos modelos avaliados em termos de RMSE e MAE para as séries temporais de Crimes contra o Patrimônio (CVP). O melhor resultado está destacado em negrito.

Abordagem	Modelo	Pernambuco		Ceará	
		RMSE	MAE	RMSE	MAE
Modelos Individuais	ARIMA	0,036	0,027	0,035	0,027
	LSTM	0,231	0,227	0,116	0,096
	MLP	0,080	0,070	0,108	0,087
	SVR	0,084	0,075	0,112	0,092
Sistemas Híbridos da literatura	ARIMA+LSTM	0,069	0,060	0,042	0,034
	ARIMA+MLP	0,034	0,025	0,034	0,027
	ARIMA+SVR	0,088	0,078	0,113	0,093
Proposta	LSTM+ARIMA	0,029	0,023	0,032	0,026
	MLP+ARIMA	0,066	0,053	0,038	0,030
	SVR+ARIMA	0,062	0,051	0,035	0,028

Conforme o ranking médio apresentado na Tabela 4, as abordagens propostas

(LSTM+ARIMA, MLP+ARIMA, SVR+ARIMA) alcançaram desempenho médio superior dentre os modelos avaliados, ocupando as três melhores posições no ranking médio. O LSTM+ARIMA ficou consistentemente entre as primeiras posições nos rankings de RMSE e MAE para todas as séries temporais analisadas, resultando na menor média de ranking (2,4). Isso evidencia que a abordagem híbrida proposta é mais robusta e eficaz na previsão das séries temporais de crimes, superando tanto os modelos individuais quanto outras combinações híbridas da literatura. Além disso, as demais versões propostas (MLP+ARIMA e SVR+ARIMA) também apresentaram desempenho superior à maioria dos modelos comparados, com médias de ranking de 3,6 e 4,2, respectivamente. Esses resultados reforçam que a integração de modelos de AM com o ARIMA contribui significativamente para a melhoria da acurácia preditiva, validando a hipótese de que a modelagem dos resíduos utilizando um modelo linear bem estabelecida pode ser um caminho promissor para modelar os padrões complexos das séries de crime.

Table 4. Ranqueamento dos modelos nas séries estudadas de acordo com as métricas RMSE e MAE.

Métodos RMSE e MAE:											
Modelo RMSE	MVI						CVP				Média
	PE		AL		CE		PE		CE		
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	
LSTM+ARIMA	1	1	5	5	4	4	1	1	1	1	2,4
MLP+ARIMA	6	6	1	2	1	1	5	5	5	5	3,7
SVR+ARIMA	2	2	2	1	10	10	4	4	3	4	4,2
MLP	3	4	4	4	2	3	7	7	7	7	4,9
ARIMA+MLP	10	10	9	9	5	5	2	2	2	2	5,6
ARIMA	9	9	7	7	6	7	3	3	3	2	5,8
SVR	3	4	3	3	8	7	8	8	8	8	6,2
LSTM	5	3	6	6	3	2	10	10	10	10	6,5
ARIMA+LSTM	8	8	8	8	9	9	6	6	6	6	7,4
ARIMA+SVR	7	7	10	10	6	6	9	9	9	9	8,2

A Tabela 5 apresenta uma análise de diferença percentual média do RMSE entre os modelos, em todas as séries temporais, onde cada célula (i, j) indica a diferença de desempenho entre o modelo i e o modelo j . Valores positivos na linha indicam uma melhoria no desempenho em termos de percentual médio do RMSE, indicando que em geral o modelo i superou o modelo j . Observa-se que o modelo LSTM+ARIMA obteve ganhos consistentes quando comparado com outros modelos, especialmente com os modelos individuais LSTM e ARIMA, reforçando a eficácia da abordagem proposta. Além disso, MLP+ARIMA também alcança desempenho destacado quando comparado com os demais modelos da literatura. No caso do SVR+ARIMA, o ganho percentual foi negativo em alguns casos devido ao desempenho alto em termos de RMSE no caso da série do MVI para o Ceará.

Table 5. Média da diferença percentual do RMSE entre os modelos avaliados. O valor médio considera todas as diferenças percentuais.

	ARIMA	LSTM	MLP	SVR	ARIMA+LSTM	ARIMA+MLP	ARIMA+SVR	LSTM+ARIMA	MLP+ARIMA	SVR+ARIMA
ARIMA	0,0	-1,4	-15,9	-0,5	13,3	-1,5	28,7	-43,8	-32,8	-10,6
LSTM	-135,2	0,0	-46,2	-34,5	-62,4	-144,1	-10,6	-195,1	-100,5	-96,2
MLP	-42,2	19,8	0,0	9,9	-10,4	-46,3	28,8	-80,3	-43,1	-35,7
SVR	-55,2	3,6	-17,0	0,0	-21,7	-60,3	18,6	-99,7	-62,2	-43,1
ARIMA+LSTM	-22,8	-5,9	-26,1	-9,8	0,0	-25,5	19,6	-71,2	-47,4	-25,2
ARIMA+MLP	1,2	-0,9	-15,7	-1,5	13,9	0,0	29,1	-42,0	-32,2	-11,9
ARIMA+SVR	-77,6	-23,5	-49,4	-33,0	-42,9	-81,8	0,0	-132,7	-96,4	-78,4
LSTM+ARIMA	27,0	34,7	23,0	32,3	38,2	25,9	52,0	0,0	8,1	21,3
MLP+ARIMA	6,7	34,5	18,3	27,2	28,1	4,2	45,3	-25,5	0,0	10,1
SVR+ARIMA	-17,5	-14,0	-30,3	1,4	3,4	-22,1	18,7	-67,0	-49,5	0,0

6. Conclusão

Este trabalho investigou o uso de sistemas híbridos para previsão de séries temporais de crimes, propondo uma abordagem que integra modelos de AM para capturar padrões não lineares e modelos estatísticos lineares para modelar os resíduos. Para validar a proposta, foram desenvolvidas três versões híbridas (LSTM+ARIMA, MLP+ARIMA e SVR+ARIMA) e comparadas com modelos individuais (ARIMA, LSTM, MLP e SVR) e com sistemas híbridos tradicionais da literatura, nos quais a modelagem linear antecede a modelagem não linear (ARIMA+LSTM, ARIMA+MLP e ARIMA+SVR). Os resultados experimentais, obtidos a partir de diferentes bases de dados criminais, evidenciaram que as versões propostas superam tanto os modelos individuais quanto as combinações tradicionais em cenários onde predominam dinâmicas complexas e não lineares. A superioridade da abordagem proposta reforça a relevância de estratégias que combinem diferentes paradigmas para lidar com a heterogeneidade dos dados criminais. Além disso, a análise dos resíduos mostrou-se fundamental para aprimorar a acurácia das previsões dos modelos de AM, indicando que a complementaridade entre modelos é um caminho promissor para desafios semelhantes em outras áreas.

Apesar dos avanços alcançados, o estudo tem limitações, como o número restrito de séries analisadas e a necessidade de explorar diferentes contextos e variáveis. Futuras pesquisas podem expandir a abordagem para outros tipos de crimes, incorporar informações exógenas, investigar novas formas de integração e seleção de modelos e utilizar novas técnicas de AM e estatísticas, visando ampliar a robustez e a generalização dos resultados.

References

- Abdulraheem, M., Awotunde, J., Oladipo, I., Adeleke, M., Ndunagu, J., Ayantola, J., and Mohammed, A. (2022). Crime rate prediction using the random forest algorithm. *LAUTECH Journal of Engineering and Technology*, 16(2):166–179.
- Alwee, R., Hj Shamsuddin, S. M., and Sallehuddin, R. (2013). Hybrid support vector regression and autoregressive integrated moving average models improved by particle swarm optimization for property crime rates forecasting with economic indicators. *The Scientific World Journal*, 2013(1):951475.
- Chainey, S. and Tompson, L. (2008). *Crime mapping case studies: practice and research*. John Wiley & Sons.
- Dakalbab, F., Talib, M. A., Waraga, O. A., Nassif, A. B., Abbas, S., and Nasir, Q. (2022). Artificial intelligence & crime prediction: A systematic literature review. *Social Sciences & Humanities Open*, 6(1):100342.
- Dave, E., Leonardo, A., Jeanice, M., and Hanafiah, N. (2021). Forecasting indonesia exports using a hybrid model arima-lstm. *Procedia Computer Science*, 179:480–487.
- de Oliveira, J. F. L., Silva, E. G., and de Mattos Neto, P. S. G. (2021). A hybrid system based on dynamic selection for time series forecasting. *IEEE Transactions on Neural Networks and Learning Systems*, 33(8):3251–3263.
- Duarte, F. C. d. L., de Mattos Neto, P. S., and Firmino, P. R. A. (2024). A hybrid recursive direct system for multi-step mortality rate forecasting. *The Journal of Supercomputing*, pages 1–34.

- Gorr, W. and Harries, R. (2003). Introduction to crime forecasting. *International Journal of Forecasting*, 19(4):551–555.
- Hajirahimi, Z. and Khashei, M. (2022). Sequence in hybridization of statistical and intelligent models in time series forecasting. *Neural Processing Letters*, 54(5):3619–3639.
- Hodson, T. O. (2022). Root mean square error (rmse) or mean absolute error (mae): When to use them or not. *Geoscientific Model Development Discussions*, 2022:1–10.
- İlgün, Esen Gül and Dener, M. (2025). Exploratory data analysis, time series analysis, crime type prediction, and trend forecasting in crime data using machine learning, deep learning, and statistical methods. *Neural Computing and Applications*. Published online: 20 March 2025.
- Izidio, D. M., de Mattos Neto, P. S., Barbosa, L., de Oliveira, J. F., Marinho, M. H. d. N., and Rissi, G. F. (2021). Evolutionary hybrid system for energy consumption forecasting for smart meters. *Energies*, 14(7):1794.
- Kang, H. W. and Kang, H. B. (2017). Prediction of crime occurrence from multi-modal data using deep learning. *PLoS ONE*, 12(4):e0176244.
- Khairuddin, A. R., Alwee, R., and Haron, H. (2020). A comparative analysis of artificial intelligence techniques in forecasting violent crime rate. In *IOP Conference Series: Materials Science and Engineering*, volume 864, page 012056. IOP Publishing.
- Khashei, M. and Hajirahimi, Z. (2019). A comparative study of series arima/mlp hybrid models for stock price forecasting. *Communications in Statistics-Simulation and Computation*, 48(9):2625–2640.
- Lab, S. P. and Hirschel, J. D. (1988). Climatological conditions and crime: The forecast is...? *Justice Quarterly*, 5(2):281–299.
- Liu, M.-D., Ding, L., and Bai, Y.-L. (2021). Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the arima to wind speed prediction. *Energy Conversion and Management*, 233:113917.
- Mandalapu, V., Elluri, L., Vyas, P., and Roy, N. (2023). Crime prediction using machine learning and deep learning: A systematic review and future directions. *IEEE Access*, 11:60153–60170.
- Manowska, A., Rybak, A., Dylong, A., and Pielot, J. (2021). Forecasting of natural gas consumption in poland based on arima-lstm hybrid model. *Energies*, 14(24):8597.
- Mao, L., Du, W., Wen, S., Li, Q., Zhang, T., and Zhong, W. (2025). Crime forecasting: A spatio-temporal analysis with deep learning models. *Journal of Computational Methods in Sciences and Engineering*, page 14727978251337993.
- Marzan, C. S., Baculo, M. J. C., de Dios Bulos, R., and Ruiz Jr, C. (2017). Time series analysis and crime pattern forecasting of city crime data. In *Proceedings of the 1st International Conference on Algorithms, Computing and Systems*, pages 113–118.
- McDowall, D., Loftin, C., and Pate, M. (2012). Seasonal cycles in crime, and their variability. *Journal of quantitative criminology*, 28:389–410.

- Muthamizharasan, M. and Ponnusamy, R. (2024). A comparative study of crime event forecasting using arima versus lstm model. *J. Theor. Appl. Inf. Technol*, 102:2162–2171.
- Noor, T. H., Almars, A. M., Alwateer, M., Almaliki, M., Gad, I., and Atlam, E.-S. (2022). Sarima: a seasonal autoregressive integrated moving average model for crime analysis in saudi arabia. *Electronics*, 11(23):3986.
- Olligschlaeger, A. M. (1997). Artificial neural networks and crime mapping. *Crime mapping and crime prevention*, 1:313.
- Pai, P.-F. and Lin, C.-S. (2005). A hybrid arima and support vector machines model in stock price forecasting. *Omega*, 33(6):497–505.
- Pastor, R. E., de Sales, J. P., Adriano Filho, F., and de Mattos Neto, P. S. (2023). Improving crime prediction through ensembles. In *Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, pages 1063–1073. SBC.
- Resende, J. P. d. and Andrade, M. V. (2011). Crime social, castigo social: desigualdade de renda e taxas de criminalidade nos grandes municípios brasileiros. *Estudos Econômicos (São Paulo)*, 41:173–195.
- Rubio, L. and Alba, K. (2022). Forecasting selected colombian shares using a hybrid arima-svr model. *Mathematics*, 10(13):2181.
- 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.
- Shah, N., Bhagat, N., and Shah, M. (2021). Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention. *Visual Computing for Industry, Biomedicine, and Art*, 4(1):9.
- Shi, J., Wang, S., Qu, P., and Shao, J. (2024). Time series prediction model using LSTM-Transformer neural network for mine water inflow. *Scientific Reports*, 14(1):18284.
- Wang, W. and Lu, Y. (2018). Analysis of the mean absolute error (mae) and the root mean square error (rmse) in assessing rounding model. In *IOP conference series: materials science and engineering*, volume 324, page 012049. IOP Publishing.
- Xu, Y., Fu, C., Kennedy, E., Jiang, S., and Owusu-Agyemang, S. (2018). The impact of street lights on spatial-temporal patterns of crime in detroit, michigan. *Cities*, 79:45–52.
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50:159–175.
- Zhang, J., Liu, H., Bai, W., and Li, X. (2024). A hybrid approach of wavelet transform, arima and lstm model for the share price index futures forecasting. *The North American Journal of Economics and Finance*, 69:102022.
- Zhang, Y., Luo, L., Yang, J., Liu, D., Kong, R., and Feng, Y. (2019). A hybrid arima-svr approach for forecasting emergency patient flow. *Journal of Ambient Intelligence and Humanized Computing*, 10:3315–3323.