

# Evaluación comparativa de algoritmos de predicción aplicados al conteo de homicidios dolosos en México

Jorge Homero García-Gómez, Sergio Ivvan Valdez,  
Hugo Carlos Martínez

Centro de Investigación en Ciencias de Información Geoespacial,  
Ciudad de México,  
México

hom.garcia@gmail.com,  
{svaldez, hcarlos}@centrogeo.edu.mx

**Resumen.** En los últimos años, los homicidios dolosos se han incrementado drásticamente en México. Esta escalada de violencia sólo ha sido analizada considerando métodos estadísticos descriptivos cuando son aplicables a todo el país, o métodos de inteligencia computacional en zonas limitadas debido a la falta de datos de alta resolución. En contraste, este trabajo usa la base de datos del Secretariado Ejecutivo de Seguridad Pública de incidencia delictiva para todos los municipios de México, para realizar una evaluación comparativa entre los modelos Autoregressive Integrated Moving Average, la red neuronal recurrente Long Short-term Memory y Neural Prophet, para la predicción del número de homicidios dolosos en los 16 municipios de Guanajuato con mayor incidencia. Los resultados muestran la diferencia de rendimiento entre los modelos mencionados, los parámetros óptimos para cada uno, los casos en los que destaca cada modelo, y permite inferir el alcance de su utilidad. La comparación se realizó mediante el error absoluto medio, el error cuadrático medio y con pruebas estadísticas de hipótesis usando bootstrap. Los modelos son competitivos entre ellos, y aunque hay una diferencia en la media del error absoluto medio y el error cuadrático medio, estos no son estadísticamente significativos, por lo que se recomienda el uso de Autoregressive Integrated Moving Average debido a su menor complejidad computacional.

**Palabras clave:** ARIMA, LSTM, neuralProphet, predicción temporal, series de tiempo, homicidios dolosos.

## Comparative Evaluation of Prediction Algorithms Applied to Counting Intentional Homicide in Mexico

**Abstract.** In recent years, intentional homicides have increased drastically in Mexico. This increase of violence has only been analyzed considering descriptive statistical methods when they are applicable to the entire country, or computational intelligence methods in limited areas due to the lack of high-resolution data. In contrast, this work uses the database of the “Secretariado Ejecutivo de Seguridad Pública” of criminal incidence for all the towns of Mexico, to

carry out a comparative evaluation between the Autoregressive Integrated Mobile Average models, the Long recurrent Short-term Memory neural network and Neural Prophet, for the prediction of the number of intentional homicides in the 16 towns of Guanajuato with higher incidence. The results show the difference in performance between the mentioned models, the optimal parameters for each one, the cases in which each model stands out, and allow inferring the scope of its usefulness. The comparison was made using the mean absolute error, the mean square error, and with statistical tests of hypotheses using bootstrap. The models are competitive among them, and although there is a difference in the mean of the mean absolute error and the mean square error, these are not statistically significant, so the use of Autoregressive Integrated Moving Average is recommended due to its lower computational complexity.

**Keywords:** ARIMA, LSTM, neuralProphet, temporal prediction, time series, intentional homicides.

## 1. Introducción

En México el número de homicidios dolosos (HD) ha tenido variaciones significativas en las últimas décadas, pues de acuerdo con [12] entre 1977 y 1992 la tendencia en los homicidios osciló entre los 17 y 21 homicidios por cada 100 000 habitantes, mientras que de 1992 hasta 2007 hubo una disminución monótona, pasando de los 20 a los 8 homicidios por cada 100 000 habitantes. Del 2007 al 2014 la tendencia osciló pasando de los 9 a los 13 homicidios por cada 100 000 habitantes [31].

Las tendencias observadas a partir del 2014 por el INEGI hasta el 2021 muestran un incremento significativo, por ejemplo en el 2020, [16] reporta 29 homicidios por cada 100 000 habitantes. Así mismo, se ha observado una disminución en el número de sentencias condenatorias para estos delitos [32]. Mientras que el aumento de la criminalidad y la ausencia del estado de derecho elevan la inestabilidad social, disminuyen la inversión privada y disminuyen el gasto público y privado, en consecuencia se reduce la actividad económica [17].

El fenómeno observado en nuestro país no es un problema que pueda atajarse o confrontarse de manera global, ya que los métodos y análisis que puedan aplicarse en algunas regiones no pueden aplicarse particularmente a otras, dada la complejidad del fenómeno [28]. Por ejemplo, el fenómeno involucra diferentes factores como son la espacialidad del territorio, la economía, la educación, la cultura, la infraestructura [28], entre otros factores que difieren en sus dinámicas y relaciones.

Otro factor que se debe considerar, se relaciona con la naturaleza, disponibilidad y calidad de los datos que los gobiernos ponen a disposición, pues estos difieren no sólo entre los países, sino que también difieren entre las jurisdicciones regionales internas de cada país [9]. Diferentes estudios que se han realizado en otras regiones han mostrado resultados convenientes no solo al caracterizar, sino al momento de combatir el fenómeno desde la realidad de cada país o ciudad.

Por ejemplo, [10] señala como las predicciones de las acciones criminales favorecen a los gobiernos, instituciones de seguridad y ciudadanía, ya que al conocer con antelación los patrones temporales delictivos, las patrullas vecinales y cuerpos de seguridad

se antepone a los momentos y lugares de alta probabilidad de que se produzca un delito. En la literatura pueden encontrarse diversos trabajos con diferentes metodologías que buscan antepone al crimen mediante la predicción de los delitos, por ejemplo, [23] utiliza el método ARIMA para el pronóstico del crimen en los municipios de Colombia, [18] es otro trabajo que para la predicción de delitos violentos en el estado de Misisipi emplea el software WEKA de código abierto, en su estudio encontraron que el algoritmo de regresión lineal mostraba el mejor desempeño.

En el caso mexicano sólo se ha encontrado un estudio de predicción de delitos que emplea datos ajenos a la capital del país, [29] realiza la predicción de los homicidios dolosos para los 15 municipios mexicanos con mayor número de casos registrados, en el estudio emplean diferentes algoritmos de aprendizaje automático, este último estudio es el más cercano al aquí realizado, sin embargo, se diferencia del mismo por los modelos y métodos de comparación empleados.

Este trabajo utiliza diferentes modelos de series de tiempo para la predicción de HD, a saber un modelo estadístico, uno de red neuronal y otro que es una combinación de estadístico y red neuronal.

Se plantea buscar los hiper-parámetros (HP) óptimos de una red neuronal recurrente de tipo LSTM (Long-short term memory, memoria a largo-corto plazo), así como la ventana de tiempo (VT) óptima del algoritmo NP (Neural Prophet, Profeta neuronal), y los parámetros óptimos del modelo ARIMA (autoregressive integrated moving average, media móvil integrada autorregresiva), para realizar el pronóstico del número de homicidios en los 16 municipios del estado de Guanajuato con mayor incidencia (en la figura 1 se muestran los municipios del estado y puede apreciarse los 16 municipios seleccionados para esta investigación), utilizando bases de datos oficiales del Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública (SESNSP), una instancia es el número de homicidios dolosos de un mes, en el periodo de tiempo comprendido entre enero del 2015 a diciembre del 2021.

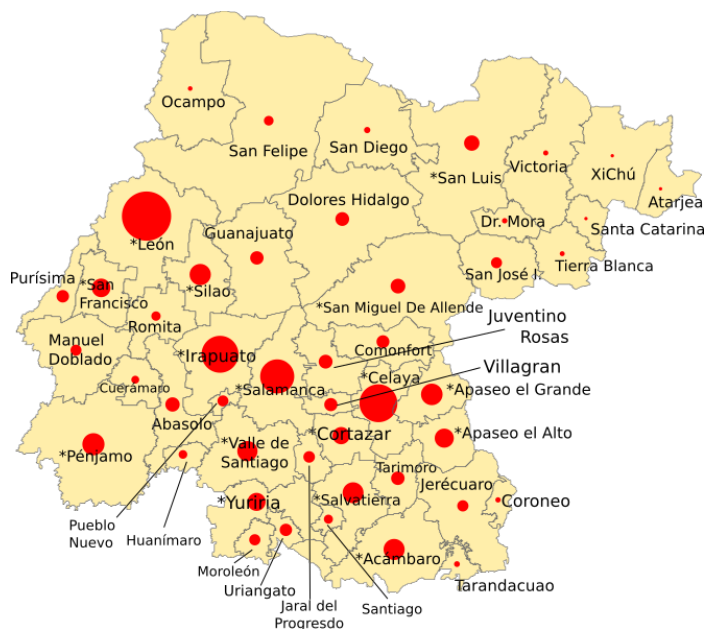
Los resultados obtenidos por los mejores modelos LSTM, NP y ARIMA son evaluados mediante la comparación de su MAE (mean absolute error, error absoluto medio), su MSE (mean squared error, error cuadrático medio) y una prueba de hipótesis bootstrap, con el objetivo de encontrar el modelo que presentó el mejor desempeño en la predicción.

En la Sección 2 se describe lacónicamente trabajos relacionados con el pronóstico de delitos. En la Sección 3 se introduce cada uno de los modelos: ARIMA, LSTM, y NP, finalmente se describe la comparación en rendimiento que se realiza entre los diferentes modelos. En la sección 4 se muestran los resultados y se discuten los mismos. Se finaliza presentando las conclusiones en la sección 5.

## **2. Trabajo relacionado**

Existen diferentes estudios que sirven para la identificación de patrones, relaciones o tendencias en los datos relacionados a la criminalidad, los estudios son principalmente de dos tipos: estudios por visualización y estudios por modelación. de acuerdo con [26], los estudios por visualización más comunes son los análisis espaciales, entre otros, estos son:

- Análisis de frecuencia de un delito en una ciudad,
- Visualización por tipo de delito por área,
- Visualización de “puntos calientes” por tipo de delito,
- Gráficos descriptivos para la identificación de relaciones:
  - Tipos de delito durante un periodo de tiempo determinado,
  - Delitos cometidos en diferentes localidades,
  - Número de delitos por hora,
  - Delitos con mayor incidencia en una ciudad.



**Fig. 1.** División política del estado de Guanajuato donde se representa, con un círculo rojo de tamaño proporcional, el número de HD acumulados del 2015 al 2021. Los 16 municipios con mayor número de HD acumulado son seleccionados para este estudio, su nombre comienza con un asterisco.

Los estudios por modelación pueden dividirse a su vez en dos subtipos, en estadísticos y de aprendizaje automático (machine learning). Uno de los modelos estadísticos usualmente empleados para el pronóstico de delitos es el modelo ARIMA, por ejemplo, [7] pronosticó los delitos en Chicago en un año utilizando los datos de los últimos 5 años con un 84 % de precisión.

Los estudios que emplean métodos de machine learning, suelen ser métodos supervisados [30], por ejemplo, para clasificación de tipos de crímenes, se utilizaron redes neuronales, máquina de soporte vectorial, k vecinos más cercanos y bosques aleatorios

[20]. Sugiriendo lo mismo que [30], quien reporta que los métodos de machine learning, más comúnmente utilizados son:

- K-Nearest Neighbours,
- Redes Neuronales,
- Boosting,
- Árboles de Decisión,
- Bosques Aleatorios.

En México se cuentan con diferentes estudios relacionados a la criminalidad que presenta el país, estudios geoestadísticos relacionados a los homicidios [13], estudios puntuales en la ciudad de México que emplean el método ARIMA para predecir delitos [21], estudios de clasificación y descubrimiento de patrones delictivos [5] o estudios que buscan optimizar rutas de patrullaje [8].

Sin embargo, la mayor parte de los estudios se concentran en la ciudad de México dado que los datos para el resto del país no son datos de alta resolución, en nuestra investigación solo se ha encontrado un estudio que emplea diferentes métodos de machine learning para el pronóstico de los homicidios que es aplicable a todo el país [29].

### 3. Metodología

La metodología seguida en esta investigación se describe en la Figura 2. Siguiendo el esquema de la Figura 2 de izquierda a derecha, se comienza por descargar los datos dispuestos por el SESNSP, que van de enero del 2015 a diciembre del 2021, el periodo seleccionado se debe a la nueva metodología de presentar los datos por parte del SESNSP, donde la información se encuentra más desagregada y específica respecto a la incidencia delictiva [14], esta nueva metodología se empezó a implementar a partir del 2015.

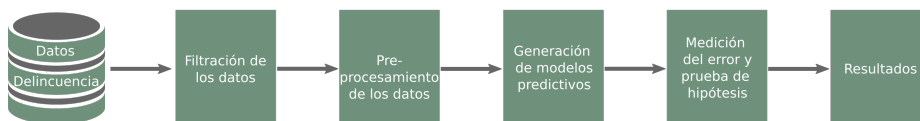


Fig. 2. Esquema general de la metodología empleada para este proyecto.

En el siguiente paso se filtran los datos, seleccionando los 16 municipios con mayor número de homicidios dolosos en el estado de Guanajuato. En el tercer paso se preprocesan los datos filtrados, transformándolos a una forma adecuada para el entrenamiento de cada modelo. Los modelos de predicción empleados son ARIMA, LSTM y NP.

Todos los modelos utilizarán 80 registros para el entrenamiento y 4 para prueba. En las siguientes subsecciones se describe cada uno de los modelos y los procedimientos seguidos en cada caso (cuarto paso en el esquema de la Figura 2), la Subsección 3.4 explica el protocolo utilizado para la evaluación y comparación de los resultados de los modelos (quinto paso en el esquema de la Figura 2).

### 3.1. Modelo ARIMA

El modelo ARIMA propuesto por [4] consiste en la combinación de dos modelos, el modelo AR (auto-regresivo) y el modelo MA (promedio móvil), la expresión del modelo ARMA se presenta en la ecuación (1):

$$y_t = - \sum_{j=1}^p \alpha_j y_{t-j} + \sum_{j=1}^q \beta_j \epsilon_{t-j}, \quad (1)$$

donde comúnmente  $\beta_0 = 1$ , mientras que el término  $y_t$  refiere al valor a predecir,  $\alpha_j$  concierne a los términos autoregresivos,  $\beta_j$  son los parámetros de las medias móviles y  $\epsilon$  es el término de perturbación estocástica, también conocido como el término de error [19]. El modelo ARIMA incluye la diferenciación ( $d$ ) de la serie con el propósito de transformarla a estacionaria, así pues, la I en la sigla ARIMA refiere a la palabra integrado (sumado) [19].

Para la obtención de los parámetros necesarios en la creación del modelo, se utilizó la biblioteca **pmdarima** desarrollada por [24]. La biblioteca **pmdarima** prueba varios conjuntos de parámetros  $p$  y  $q$ , seleccionando el modelo que minimiza el criterio de Información de Akaike (AIC) propuesto por [2]. Mientras que para la selección del término de diferenciación  $d$ , utiliza la prueba de Dickey-Fuller aumentada de [22].

El procedimiento para la obtención del modelo ARIMA para cada uno de los 16 municipios seleccionados, inició por la conversión de los datos, a una estructura de datos de serie de tiempo, posteriormente se seleccionaron los parámetros  $p$ ,  $d$ ,  $q$ , y se realizó el ajuste del modelo a cada una de las series de tiempo. A continuación se realizó la predicción del modelo, y finalmente se almacenaron los resultados de cada municipio.

### 3.2. Modelo LSTM

El algoritmo LSTM propuesto por [15] se caracteriza por no presentar el problema de las estructuras RNN conocidos como “desaparición del gradiente”, especialmente observados en secuencias de entradas grandes [1].

Además presentan un mejor rendimiento de convergencia [11] y son convenientes (al igual que RNN) al trabajar con bases de datos con componente temporal o secuencial. Las redes LSTM se componen de múltiples funciones en comparación de una RNN, dichas funciones pretenden olvidar los datos innecesarios de las entradas y procuran recordar o conservar los datos de utilidad a la predicción. Las funciones que

integran una red LSTM se presentan a continuación:

$$f_t = g(W_f y_t + U_f h_{t-1} + b_f), \quad (2)$$

$$k_t = \tanh(W_k y_t + U_k h_{t-1} + b_k), \quad (3)$$

$$c_t = f_t c_{t-1} + i_t k_t, \quad (4)$$

$$o_t = g(W_o y_t + U_o h_{t-1} + b_o), \quad (5)$$

$$h_t = o_t \tanh(c_t), \quad (6)$$

donde  $y_t$  es la entrada de datos en el tiempo  $t$  y  $g$  es una función de activación no lineal como la función sigmoide o *ReLU*.  $W$ ,  $U$  son matrices de ponderaciones y  $b$  es el vector conocido como sesgo.  $h_t$  y  $c_t$  son la salida y vector de estado de celda en el tiempo  $t$ .  $f_t$  es utilizado para recordar información antigua y sirve para obtener información nueva [1].

Para la predicción se utiliza el concepto de ventana móvil temporal (*VT*), el cual consiste en utilizar  $h$  pasos de tiempo y posteriormente ir desplazando  $i$  pasos la ventana temporal. En este caso el objetivo es predecir un mes a la vez, por lo que la ventana a utilizar queda determinada como  $k = t - h, t - h + 1, \dots, t - 1$ .

Para este experimento se realizaron las predicciones utilizando únicamente la variable endógena del municipio a predecir (al igual que en *ARIMA*). Los hiper-parámetros de los modelos LSTM se buscaron exhaustivamente en el conjunto de las combinaciones siguientes:

- Se probaron VT desde 2 hasta 12 meses,
- Se probó dropout en 0 y 0.2,
- Se probaron las funciones de activación sigmoide y tanh,
- Se probaron números de neuronas de 20, 40 y 60.

Las diferentes combinaciones producen 132 resultados en total. Todos los modelos se desarrollaron con un arreglo de dos capas donde la segunda capa tenía la mitad del número de neuronas de la primera capa. Antes de aplicar la LSTM se les quito la tendencia a los datos restandoles un modelo de regresión lineal, y se normalizaron, restando el promedio y dividiendo entre la desviación estándar.

### 3.3. Neural Prophet

NP es un modelo reciente propuesto en noviembre de 2021 por [27], y es una extensión del algoritmo Facebook Prophet propuesto en 2018 por [25]. El modelo NP esta integrado por seis componentes tal como se muestra en la siguiente ecuación:

$$Y_t = T(t) + S(t) + E(t) + F(t) + A(t) + L(t), \quad (7)$$

donde  $T(t)$  corresponde a la tendencia de la serie temporal después de eliminar la estacionalidad,  $S(t)$  corresponde a la estacionalidad y  $E(t)$  corresponde a los eventos “especiales” preestablecidos. Estas primeras tres componentes se corresponden a la parte estadística del algoritmo NP, cada una de sus componentes se ajusta individualmente.

Las componentes  $L(t)$  son las covariables actuales retrasadas  $n$  periodos de tiempo, mientras que el pronóstico para estas covariables se modela con el término  $F(t)$ . La componente  $A(t)$  conocida como término autorregresivo aplica el concepto de observación de valores pasados, y usa los mismos para la predicción de valores futuros.

Las ultimas tres componentes del modelo NP se corresponden a la parte de red neuronal, por lo que NP es un algoritmo con una parte estadística y una parte de red neuronal, de acuerdo con [6] esta combinación presenta la ventaja de proporcionar una ganancia en eficiencia al tiempo que se limita la pérdida de interpretabilidad.

Las pruebas realizadas con NP se realizaron de manera análoga a las realizadas con el modelo LSTM, utilizando las variables endógenas correspondientes al municipio que se predice. Mencionar que en cuanto al  $VT$  se realizó el mismo número de combinaciones que en los modelos LSTM, por que que se generaron en este caso 11 resultados con este modelo.

### 3.4. Evaluación de modelos

Una vez obtenidas las predicciones por los diferentes modelos, se evalúa su desempeño empleando el MAE y MSE, y una prueba de hipótesis de medias, para la media del MAE, usando el método bootstrap.

El método bootstrap, sirve, entre otras cosas, para hacer pruebas de hipótesis [3] no-paramétricas, y consiste en tomar repetidas muestras de una población con reemplazo, estas muestras son observaciones de la distribución empírica, con ellas se pueden calcular estimadores que se aproximan a los de la distribución subyacente de los datos. En este caso estimamos la probabilidad de que la media del MAE producido por un algoritmo sea menor que la media de otro, el resultado es un estimador del  $p$  – valor de la prueba de hipótesis de medias.

Para comparar los modelos se calcula el MAE y MSE sobre los datos de prueba para los 16 municipios con mayor número de homicidios del estado de Guanajuato. Todos usan la misma cantidad de datos de prueba, que son 4 por municipio, 80 instancias de entrenamiento.

Ya que el número de combinaciones de parámetros que se probaron es exhaustiva, se presenta los resultados para el modelo ARIMA con la mayor AIC, usando para cada municipio los valores de  $d$ ,  $p$  y  $q$  que maximizan el AICc. También se comparan los 3 modelos NP y los 11 de LSTM con el MAE menor. Finalmente, se realizaron pruebas de hipótesis que comparan las medias de cada modelo contra todos los demás, de la siguiente forma:

- Hipótesis nula ( $H_0$ ): Indica que el rendimiento de un modelo A es igual a un modelo B,
- Hipótesis alternativa ( $H_1$ ): Indica que la media del error del modelo A es menor que la media del error del modelo B.

## 4. Resultados y discusión

El objetivo de esta sección es comparar el desempeño en el MAE y MSE de los diferentes algoritmos así como reportar los parámetros óptimos, dentro del conjunto de



búsqueda, que se encontraron mediante pruebas exhaustivas. El Cuadro 1 muestra los modelos ordenados por MAE.

**Tabla 1.** Resultados de los modelos con mejor desempeño, seleccionados para compararlos con la prueba de hipótesis donde ninguno resultó significativamente diferente. Se muestran los hiperparámetros del modelo y su respectivo MAE y MSE.

				MAE	MSE
ARIMA				3.210557	21.003021
<b>NP</b>					
VT	parámetros				
11				3.258145	26.683529
2	por omisión			3.297269	24.408899
3				3.44949	25.374554
<b>LSTM</b>					
VT	nn	$f_{\text{activación}}$	dropout		
8	60	tangh	0	3.580727	24.947718
7	40	tangh	0.2	3.602632	25.467693
8	20	tangh	0.2	3.61437	23.493635
12	60	tangh	0.2	3.629066	25.112785
2	40	sigmoid	0.2	3.641926	23.37434
11	40	sigmoid	0.2	3.665473	26.507914
4	60	sigmoid	0.2	3.681434	26.659206
6	20	sigmoid	0.2	3.692786	27.279583
5	40	tangh	0.2	3.706302	26.986547
3	20	sigmoid	0.2	3.724117	26.862549
2	60	tangh	0.2	3.733833	24.569363

Usualmente se observa que existe una correspondencia entre el MAE y el MSE, sin embargo, se advierte en el Cuadro 1 que algunos modelos arrojan un mejor resultado en el MAE mientras que otros en el MSE, esto se debe a que algunos errores están entre 0 y 1, y cuando se elevan al cuadrado se hacen menores.

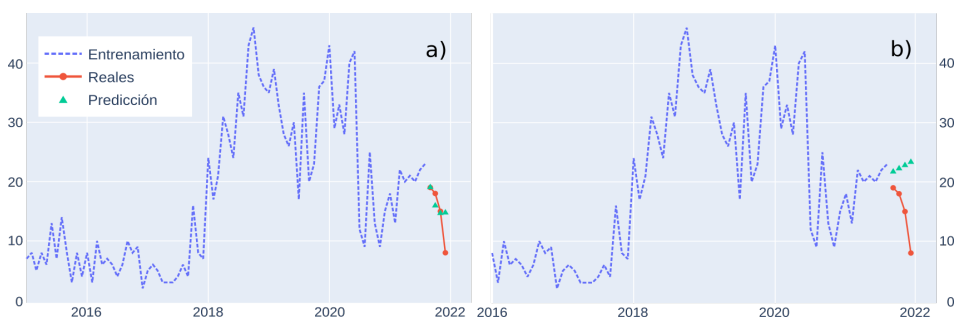
De los resultados puede observarse, en el Cuadro 1, que el modelo ARIMA arroja el mejor resultado, sin embargo, la prueba de hipótesis no permitió rechazar la hipótesis nula, esto es, se confirmó que no existe diferencia estadística entre los modelos. Los desempeños observados en los modelos LSTM y NP, puede deberse a que los registros dispuestos por SESNSP son insuficientes para el correcto entrenamiento de los modelos, pues estos modelos fueron ideados para la predicción de grandes conjuntos de datos.

Otro punto es que para ARIMA los parámetros  $q$ ,  $p$  y  $d$  se buscan para cada municipio, mientras que los errores reportados para LSTM y NP usan los mismos hiper-parámetros para todos los municipios. Finalmente, una búsqueda en un conjunto mayor de valores y que incluya más hiperparámetros podría mejorar los resultados, pero dados los resultados actuales, no se espera una gran mejora.

La Figura 3 muestra las predicciones para el municipio de Irapuato con ARIMA en la izquierda y LSTM a la derecha. Es evidente que ARIMA tiene una predicción con

menos error para los 4 meses de prueba, sin embargo, LSTM parece seguir la tendencia de los últimos meses para la predicción.

Por lo tanto, LSTM no puede capturar el patrón de la red, se puede deber a hiper-parámetros incorrectos, por ejemplo, el VT de 12 puede ser el mejor en promedio, pero no el mejor para el caso específico de Irapuato, otra razón puede ser que la cantidad de datos es insuficiente, y por último, que existen factores del modelo como otros valores de hiper-parámetros o normalización que pueden mejorar los resultados, esta última se considera la menos probable, ya que se realizó una búsqueda exhaustiva en ciertas combinaciones de hiper-parámetros, por lo que una búsqueda más fina se espera que sólo mejore marginalmente, en este sentido, la vía más admisible sería buscar entre hiper-parámetros que no se optimizaron en este trabajo u otros factores como la estandarización o normalización de los datos.



**Fig. 3.** La imagen de la izquierda (a) muestra los resultados del modelo ARIMA en la predicción de los HD en el municipio de Irapuato, mientras que la imagen de la derecha (b) muestra los mismos resultados obtenidos con el modelo LSTM con una VT de 12, 60 neuronas en la primera capa y sin dropout. El conjunto de entrenamiento se representa en color azul, las líneas rojas son los valores reales, mientras que la línea verde muestra las predicciones realizadas por los modelos.

## 5. Conclusiones

Este trabajo explora tres modelos, ARIMA, LSTM y Neural Prophet para la predicción mensual del número de homicidios dolosos en los 16 municipios con mayor incidencia del estado de Guanajuato, México. Los resultados obtenidos muestran que el modelo ARIMA presenta el menor MAE y MSE, sin embargo, se encontró que no existe una diferencia estadística significativa entre los modelos, Por lo que se recomienda el empleo del modelo con menor complejidad y coste computacional, el modelo ARIMA.

En contraste con los resultados encontrados, hay un gran auge en el uso de LSTM y Neural Prophet para la predicción de series de tiempo, por lo que se considera que aún se deben explorar modificaciones a los experimentos presentados en este estudio con el fin de intentar mejorar su desempeño. A partir de los resultados obtenidos, se consideran 3 posibles explicaciones por las que ARIMA resultó con mejor desempeño:

1. Que la cantidad de datos para entrenar los modelos LSTM y NP es insuficiente.

2. Que se pueden buscar hiper-parámetros y otras variables como normalización o estandarización diferente de los datos, e información de los municipios adyacentes.
3. Que cada municipio requiere hiper-parámetros diferentes, ARIMA utiliza un conjunto  $\{p, q, d\}$  para cada municipio, mientras que el modelo de LSTM y NP se reporta con los mismos hiperparámetros para todos los municipios.

A partir de estos argumentos, el trabajo futuro explorará :

- Incorporar los datos anteriores a la nueva metodología del SESNSP, lo que incrementaría en cuatro años los datos,
- Incluir en los modelos LSTM y NP las serie temporal concerniente al municipio que se desea realizar la predicción y la serie del municipio adyacente con mayor número de HD,
- Reportar los errores de predicción por municipio. Aunque una de las razones para no hacerlo así es que la cantidad de datos es poca, y un estadístico sobre 4 datos de prueba no es suficiente para inferir diferencias estadísticas.

## Referencias

1. Aggarwal, C. C.: Neural networks and deep learning. Springer, vol. 10, pp. 419–458 (2018) doi: 10.1007/978-3-319-94463-0
2. Akaike, H.: Stochastic theory of minimal realization. IEEE Transactions on Automatic Control, vol. 19, no. 6, pp. 667–674 (1974) doi: 10.1109/TAC.1974.1100707
3. Bartlett, J.: The miracle of the bootstrap The Stats Geek
4. Box, G., Jenkins, G.: Time series analysis: forecasting and control. Water Science and Technology, (1976)
5. Campedelli, G. M.: Explainable machine learning for predicting homicide clearance in the United States. Journal of criminal justice, vol. 79 (2022) doi: 10.1016/j.jcrimjus.2022.101898
6. Catherine, V.: In-depth understanding of NeuralProphet through a complete example. Towards data science (2022)
7. Cesario, E., Catlett, C., Talia, D.: Forecasting crimes using autoregressive models. In: IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd International Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress. pp. 795–802. IEEE (2016) doi: 10.1109/DASC-PICOM-DataCom-CyberSciTec.2016.138
8. Chainey, S., Matias, J., Junior, F., Silva, T., Macedo, J., Pires Magalhaes, R., de Queiroz Neto, J. F., Silva, W.: Improving the creation of hot spot policing patrol routes: Comparing cognitive heuristic performance to an automated spatial computation approach. ISPRS International Journal of Geo-Information, vol. 10, no. 8, pp. 560 (2021) doi: 10.3390/ijgi10080560
9. Dávila, E. A. S.: Rentabilidad de los recursos que la políticas públicas municipales ponen a disposición de las PYME'S considerando su cultura organizacional. Academia Journals, vol. 8, no. 3, pp. 1882–1887 (2016)
10. Devarakonda, D. S.: Time series analysis and forecasting of crime data. Ph.D. thesis, Master Dissertation Thesis, California State University (2019)
11. DiPietro, R., Hager, G. D.: Deep learning: RNNs and LSTM. In: Handbook of medical image computing and computer assisted intervention, pp. 503–519. Elsevier (2020)

12. Escalante, G. F.: Panorama del homicidio en México. Esquema de análisis territorial 1990-2007. Los grandes problemas de México. Seguridad nacional y seguridad interior, pp. 301–330 (2010)
13. Fuerte-Celis, M. P., Sánchez-Castañeda, M. D.: Clusters of violence in Mexico: An analysis of homicide rates from 2000 to 2012. *Journal of Latin American Geography*, vol. 20, no. 1, pp. 99–133 (2021)
14. Gobierno de México: Incidencia delictiva del secretariado ejecutivo del sistema nacional de seguridad pública (2022)
15. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural computation*, vol. 9, no. 8, pp. 1735–1780 (1997) doi: 10.1162/neco.1997.9.8.1735
16. INEGI: Datos preliminares revelan que en 2020 se registraron 36 579 homicidios (2021)
17. Loría, E.: Impacto de secuestros y homicidios en la inversión extranjera directa en México. *Contaduría y administración*, vol. 65, no. 3 (2020) doi: 10.22201/fca.24488410e.2020.2246
18. McClendon, L., Meghanathan, N.: Using machine learning algorithms to analyze crime data. *Machine Learning and Applications: An International Journal (MLAIJ)*, vol. 2, no. 1, pp. 1–12 (2015) doi: 10.5121/mlaj.2015.2101
19. Montenegro, A.: Análisis de series de tiempo. Bogotá: Pontificia Universidad Javeriana, (2010)
20. Mosquera Cabra, J.: Trabajo de grado sobre predicción de los tipos de delitos en Medellín. Tesis de Licenciatura, Universidad de Antioquia, Facultad de Ciencias Económicas, (2021)
21. Pambabay-Calero, J., Bauz-Olvera, S., Flores-González, R., Piña-García, C.: Multivariate analysis and characterization of low impact crime in Mexico City. *F1000Research*, vol. 10, no. 1299, pp. 1–22 (2021) doi: 10.12688/f1000research.72990.1
22. Phillips, P. C., Perron, P.: Testing for a unit root in time series regression. *Biometrika*, vol. 75, no. 2, pp. 335–346 (1988)
23. Santos-Marquez, F.: Spatial beta-convergence forecasting models: Evidence from municipal homicide rates in Colombia. *Journal of Forecasting*, vol. 41, no. 2, pp. 294–302 (2022)
24. Smith, T. G.: ARIMA estimators for Python, about the project (2017)
25. Taylor, S. J., Letham, B.: Forecasting at scale. *The American Statistician*, vol. 72, no. 1, pp. 37–45 (2018) doi: 10.1080/00031305.2017.1380080
26. ToppiReddy, H. K. R., Saini, B., Mahajan, G.: Crime prediction & monitoring framework based on spatial analysis. *Procedia computer science*, vol. 132, pp. 696–705 (2018) doi: 10.1016/j.procs.2018.05.075
27. Triebe, O., Hewamalage, H., Pilyugina, P., Laptev, N., Bergmeir, C., Rajagopal, R.: NeuralProphet: Explainable forecasting at scale. *arXiv preprint arXiv:2111.15397*, (2021) doi: 10.48550/arXiv.2111.15397
28. Valasik, M., Brault, E. E., Martinez, S. M.: Forecasting homicide in the red stick: Risk terrain modeling and the spatial influence of urban blight on lethal violence in Baton Rouge, Louisiana. *Social Science Research*, vol. 80, pp. 186–201 (2019) doi: 10.1016/j.ssresearch.2018.12.023
29. Valdez, S. I., Hernández-Baena, A.: On the best-performed time window size for homicide count forecasting. In: 2021 Mexican International Conference on Computer Science (ENC). pp. 1–8. IEEE (2021) doi: 10.1109/ENC53357.2021.9534830
30. Vaquero-Barnadas, M.: Machine learning applied to crime prediction. B.S. thesis, Universitat Politècnica de Catalunya (2016)
31. Zepeda Lecuona, G. R., Jiménez Rodríguez, P.: Impunidad frente al homicidio doloso en México. *Este País*, vol. 308, pp. 13–22 (2016)
32. Zepeda Lecuona, G. R., Jiménez Rodríguez, P.: Impunidad en homicidio doloso en México: reporte 2019. *Impunidad Cero*, (2019)