

Análisis Criminal Inteligente Usando Eventos Geográficos y Temporales con la Red Neuronal de Máxima Sensibilidad

Intelligent Crime Analysis Using Geographical, and Temporal Events by Maximum Sensibility Neural Network

Torres Francisco, Neira, Leticia & Torres-Treviño, Luis

Resumen. En este artículo se propone la predicción de eventos criminales usando datos históricos. Se propone una arquitectura usando una red neuronal de máxima sensibilidad considerando patrones temporales y geográficos con una serie de delitos criminales más frecuentes. El sistema puede ser usado para realizar predicciones sobre eventos futuros y establecer posibles causas, relación entre eventos similares, puntos calientes “hot spots” de delitos específicos, etc. El sistema fue probado considerando datos históricos de tres años de una ciudad mexicana (2011-2013) y los resultados ilustran la utilidad del análisis de estos sistemas para predecir ciertos eventos.

Palabras Claves. Red Neuronal, Crímenes Comunes, Sistema para toma de decisiones

Abstract. Prediction of crime events using historical data is proposed in this paper. An architecture is proposed using a maximum sensibility neural network considering as inputs temporal, and geographical events with a set of common crime. The system can be used to make predictions on future events establishing possible causes, relations between similar events, hot spots for specific crimes, etc. The system was tested considering historical records of three years from a Mexican city (2011-2013) and results illustrate usefulness of the analysis that can be made with this system for crime events predictions.

Keywords. Neural Networks, Common Crime, Support Decision System

Introducción

Para el análisis de datos con eventos criminales existen diversas herramientas de las que podemos destacar las basadas en reglas de asociación o clasificación, las basadas en técnicas de regresión o estadísticas y las basadas en redes neuronales. Ejemplos de artículo basados en reglas de asociación o de agrupamiento u otras técnicas que son muy utilizadas en la minería de datos, tenemos los siguientes trabajos, primero en el trabajo de Tayal, et al. Integraron diversos módulos donde e destaca el método de *K-means* para la detección e identificación de delitos así como su predicción [1].

En [2], se presenta un sistema de predicción que se basa en la información que proviene de Twitter y por medio de un análisis lingüístico y estadístico se estima la densidad de una categoría que se asocia a un delito en particular por lo que es útil para realizar predicciones de delitos en zonas y tiempos específicos.

En el trabajo de Ku y Leroy, desarrollan un sistema de apoyo en la toma de decisiones basado en un clasificador Naive Bayes para clasificar diferentes delitos considerando información que proviene del análisis de textos de reportes criminales [3]. Park propone el uso de cadenas de Markov para predecir la existencia de un delito futuro considerando la probabilidad de que ocurra dicho delito; de esta manera, en cada cuadrante se predice la probabilidad de ocurrencia de cada uno de los delitos [4]. Se han realizado diferentes comparaciones entre

diversas técnicas considerando datos sobre diversos delitos. En el trabajo de Oatley y Ewart [5] se realiza una comparación entre diversas técnicas como teoría de redes sociales, epistemología, autómatas celulares, redes Bayesianas y estadística espacial. El objetivo es desarrollar un sistema de apoyo en la toma de decisiones para relacionar delitos con perfiles de los delincuentes, identificar redes criminales, generar sospechosos, predecir actividades criminales y ligar crímenes. Una historia breve del análisis de datos criminales se realiza en el trabajo de Shane Johnson [6] donde se revisa el impacto que tiene las herramientas de análisis espacial y temporal para la predicción y el desarrollo de modelos de localización de diversos crímenes. Usando información geográfica sobre sucesos criminales y por medio de una red neuronal basada en un mapa de auto-organización, se realiza minería de texto sobre reportes criminales narrativos que no están estructurados para explorar información oculta que relacione patrones geográficos con la atención de los delito [7].

Dentro de las técnicas de redes neuronales existen una variedad de ellas y cada una ofrece ventajas y desventajas. La red neuronal de Perceptrón multicapas es la más utilizada porque genera modelos muy precisos, sin embargo requiere de muchos datos y de mucho tiempo de entrenamiento lo cual la limita para aplicaciones en tiempo real. Otras redes que se basan en métodos iterativos son la de base radial y la basada en vectores de soporte.

Otro tipo de redes neuronales se basan en la auto-organización de la información considerando distancias entre patrones previamente aprendidos. Ejemplo de estas redes se encuentra la red neuronal con memorias de auto-organización, la red neuronal de máxima sensibilidad y la red neuronal basada en teoría de resonancia asociativa.

En [8] se desarrolla un sistema para el apoyo de la toma de decisiones basado en una red neuro-difusa de un mapa de auto-organización para detectar y analizar patrones de tendencias de actividades criminales. También se incluye un mecanismo de extracción de reglas lingüísticas que pueden ser útiles para realizar estrategias de seguridad y sobre la aplicación de las leyes.

En [9] se propone una clasificación de minería de datos criminales con puntos geográficos críticos que está basada en una red neuronal con vectores de soporte. Por medio de este sistema se predice la localización geográfica de determinados delitos.

Para el pronóstico de delitos se pueden aplicar técnicas de regresión lineal, sin embargo, en muchos casos existen relaciones no lineales entre los datos, por lo tanto en [10] se propone un sistema híbrido donde se combina una red neuronal basado en vectores de soporte y en herramientas de autoregresión (ARIMA) además se utiliza un algoritmo de optimización basado en cúmulos de partículas para ajustar los parámetros de dichos modelos.

El presente trabajo presenta el uso de la red neuronal de máxima sensibilidad que es muy similar a una red de auto-organización con un mecanismo dinámico de aprendizaje, por lo cual puede aprender en tiempo real e intercambiar su funcionalidad de funcionamiento y de aprendizaje.

Para el análisis de datos criminales se han utilizado diversas herramientas que van desde las basadas en reglas de agrupamiento, en formulaciones probabilísticas y en relaciones

funcionales por medio de sistemas de regresión lineal o con redes neuronales. Existen diversas redes neuronales (más de 50) sin embargo, las que se han probado con mucha eficiencia son las redes de multicapas llamadas Perceptrón, las de base radial y las que contienen memorias de auto-organización.

En el presente trabajo se requiere establecer una relación funcional entre las posiciones geográficas donde suceden delitos del fuero común, el mes, día, hora y minuto con cada uno de los delitos tipificados. En teoría cualquier herramienta puede ser utilizada para generar, por ejemplo predicciones sobre eventos futuros; sin embargo, en la práctica cuando se compromete la eficiencia con la rapidez de respuesta, no cualquier herramienta es útil.

La desventaja que tienen las herramientas de regresión es que requieren ciertas condiciones en los datos para evitar problemas de singularidad que impiden calcular los regresores; aún si se resuelve los datos pueden estar fuertemente correlacionados con relaciones no lineales, lo cual al generar los métodos basados en regresión puede tener falta de ajuste y generar un error grande en la predicción.

Las redes neuronales son las que más se han utilizado para generar modelos de relación. La red neurona de Perceptrón multicapas es la más utilizada, sin embargo requiere una cantidad significativa de datos y mucho tiempo de procesamiento para ajustar los pesos de la red que son los que permiten su funcionamiento. Este proceso es iterativo lo cual incrementa significativamente el tiempo de procesamiento que puede ser prohibitivo si es necesario utilizar la herramienta en un momento dado, es necesario esperar hasta que termine el entrenamiento antes de pensar utilizar la red. Una vez entrenada si surgen nuevos datos, no es posible incorporarlos al aprendizaje ya que es necesario repetir todo el procedimiento de entrenamiento con los nuevos registros.

Una de las redes que se han utilizado más en el análisis de datos criminales son las basadas en memorias de auto-organización. La red neuronal de máxima sensibilidad pertenece a esta clase donde existen mecanismos autónomos de organización por lo que no dependen de procesos iterativos y por lo tanto, es posible entrenar la red con una cantidad limitada de patrones (desde uno) y cambiar al procedimiento de funcionamiento para usar la red. En el momento que surgen nuevos patrones, es posible regresar al proceso de entrenamiento inmediatamente para incorporar los nuevos patrones y volver a la etapa de funcionamiento en un instante.

La red neuronal de máxima sensibilidad

La red neuronal de máxima sensibilidad se clasifica como una de auto organización. Esta red tiene mecanismos de aprendizaje muy sencillos y que no son iterativos, por lo que pueden ser utilizadas para el aprendizaje en línea. De esta manera es la red neuronal ideal para poder realizar el aprendizaje y el uso de la red en cualquier instante.

La red neurona de máxima sensibilidad tiene la arquitectura mostrada en la Figura 1. La red neuronal se basa en el principio de la máxima sensibilidad donde se presenta un patrón a las entradas de la red y todas las neuronas son activadas con diferentes grados. Por medio de un mecanismo del tipo “el ganador se lleva todo” se establece solamente una neurona ganadora

y si su nivel de activación supera una señal de umbral preestablecido, entonces la neurona de actuación asociada se activa y su señal se presenta a la salida. En caso contrario, todas las neuronas contribuyen con sus señales de activación y sus neuronas de actuación para generar una salida generalizada.

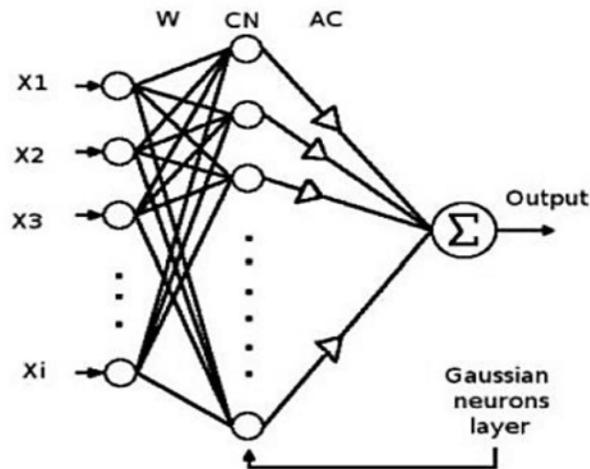


Figura 1. Arquitectura de la red neuronal de máxima sensibilidad

El aprendizaje está gobernado por la misma propiedad de máxima sensibilidad. Si hay máxima sensibilidad, los pesos de la neurona ganadora incluyendo la neurona de actuación asociada son ajustados considerando el nuevo patrón para aproximarlos al centro de masa. Cuando no hay máxima sensibilidad es integrada una nueva neurona del total disponible para guardar el patrón. Cuando se terminan las neuronas del total disponible entonces se busca la neurona que ha sido poco utilizada (no ha ganado) para reemplazar sus pesos que representan a patrones poco o nada utilizados por el nuevo. Todas las neuronas tienen asociado un peso de uso que va en decremento con un factor de olvido constante, por lo que cada vez que se utiliza la red neuronal, pero si existe máxima sensibilidad entonces este peso de uso se actualiza al valor inicial por lo que su posibilidad de ser borrado disminuye. Este mecanismo es muy útil para adaptar la red y adaptarse a las nuevas tendencias que presenten los patrones y desechar patrones que son poco recurrentes o inclusive que sean producto del ruido. La desventaja que se tiene es que la red neuronal se especializa al comportamiento del fenómeno y si es utilizada en otro contexto, su capacidad de generalización no es útil.

Sistema propuesto para la predicción de eventos criminales

El sistema propuesto requiere de un modelo para predecir eventos delictivos en tiempo real y que se ajuste a las tendencias que vayan existiendo en la zona. Aunque en un principio se basa en la información histórica, poco a poco la red neuronal va incluyendo nuevos patrones y se especializa a las condiciones del entorno.

El sistema de predicción relaciona las condiciones de entorno (el clima), la temporalidad (mes, día, hora, minuto) y la posición geográfica con los diferentes delitos del fuero común. (Figura 2) Este tipo de relaciones funcionales las puede realizar en tiempo real y es posible incorporar nueva información sin esperar una etapa prolongada de “entrenamiento” o “aprendizaje” donde no se puede utilizar la herramienta, por lo que es posible usar la herramienta y cambiar al entrenamiento en un instante.

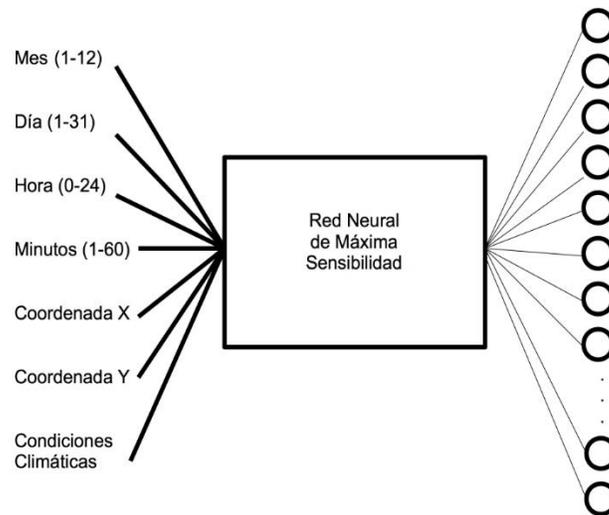


Figura 2. Arquitectura utilizada por el sistema predictivo

El mecanismo de aprendizaje de la red neuronal de máxima sensibilidad, le permite a la arquitectura propuesta un aprendizaje muy rápido, por lo que si se tiene información histórica correcta y suficiente (al menos tres años) se tendrá una buena eficiencia en el desempeño. Es importante destacar que la red tiene a especializarse por lo que puede modelar el efecto que pueden tener ciertas políticas o leyes sobre la ocurrencia de ciertos delitos. En otras palabras, la herramienta adquiere una “personalidad” adecuada a las situaciones de la región donde se utiliza o se alimenta la red.

El sistema de modelado se basa en eventos temporales, geográficos y ambientales por lo que puede ser utilizada para otras aplicaciones diferentes a las de la predicción, tales como el agrupamiento de zonas de conflicto, identificación de patrones de comportamiento, simulación de eventos etc., que pueden ser útiles para una mejor toma de decisiones.

Demostración de funcionamiento de la herramienta

La arquitectura consiste en una red de 6 entradas donde se considera el mes, el día del mes, la hora en formato de 24 horas y los minutos de 1 a 60, las coordenadas (x, y) donde sucedió el delito y se relaciona con cada uno de los delitos tipificados del fuero común. Si se considera el clima se establecerá un aumento de dos entradas más para la temperatura y la humedad.

En la actualidad la arquitectura de la red se consideran 16 salidas donde marca la existencia de un delito específico. La tabla 1 muestra cada uno de estos delitos tipificados. Se les asigna un número entero para establecer la salida que se activará con un grado de fuerza dependiendo de la posibilidad que un delito se presente. Como cada delito puede presentarse con condiciones temporales y geográficas similares. Para utilizar la red neuronal se establece un momento temporal considerando un mes, un día, una hora y uno minuto específico y una posición geográfica, se predicen los posibles delitos que se pueden presentarse en las condiciones ya mencionadas.

Tabla 1. Delitos del fuero común

Delito
Robo a banco (Con violencia)
Robo a banco (sin violencia)
Robo a casa habitación (Con violencia)
Robo a casa habitación (sin violencia)
Robo a negocio (Con violencia)
Robo a negocio (sin violencia)
Robo a persona (Con violencia)
Robo a persona (Sin violencia)
Robo de vehículo totalidad (Con violencia)
Robo de vehículo totalidad (Sin violencia)

Considerando la siguiente entrada siendo muy específicos en el tiempo en una posición geográfica específica.

Tabla 2. Entrada Fecha y Hora

Mes	día	hora	minuto
5	15	22	00

Tabla 3. Salida las siguientes ponderaciones de los delitos

Peso Posibilidad de Ocurrencia	Delitos
0.0000008	Robo a banco con violencia
1.754D-08	Robo a banco sin violencia
0.0000016	Robo a casa habitación con violencia
0.5295351	Robo a casa habitación sin violencia

0.1188201	Robo a negocios con violencia
0.0226574	Robo a negocios sin violencia
0.0901080	Robo a personas con violencia
0.0292420	Robo a personas sin violencia
0.0588343	Robo de vehículos (totalidad) con violencia
0.2651825	Robo de vehículos (totalidad) sin violencia

En primer lugar tenemos el robo a casa habitación sin violencia. En segundo lugar tenemos el robo de vehículos. En tercer lugar tenemos el robo a negocios con violencia y en cuarto lugar tenemos el robo a personas con violencia. Para estos delitos buscamos fechas, horas y posiciones geográficas similares para verificar si se han presentado dichos delitos.

Tenemos los siguientes registros del año 2013 que son utilizados para verificar las respuestas de la predicción, recordemos que la predicción no es una probabilidad de ocurrencia, sino más bien una ponderación sobre la posibilidad de que un delito ocurra considerando la información histórica:

Tabla 4. Robo casa habitación sin violencia

Mes	día	hora	minuto
5	15	19	57
5	15	22	14
5	15	22	0

Tabla 5. Valores que se presenta en el mismo cuadrante, mismo día a una diferencia de horas también tenemos los siguientes registros de robo de vehículos sin violencia:

Mes	día	hora	minuto
5	15	18	34
5	15	23	17
5	15	21	15

Tabla 6. También tenemos los robos a negocios con violencia

Mes	día	hora	minuto
5	15	16	36

Finalmente tenemos los datos de delitos por robo a personas:

Tabla 7. Robo a personas

Mes	día	hora	minuto
5	15	18	11

Sin embargo ya se encuentran lejanos de la hora solicitada así como lejana de la posición geográfica. Conforme se alejan más los eventos delictivos, las ponderaciones se reducen drásticamente.

Conclusiones

La predicción de delitos como otros sistemas de predicción no son sencillos y la eficiencia del predictor puede estar sujeto a infinidad de circunstancias, cambios de políticas de seguridad, la instalación de equipos de vigilancia, nuevos parques, hospitales, comandancias, etc. todos estos factores influyen para que los eventos delictivos cambien. Al realizar un modelo basado en una red neuronal de tipo multicapas o un modelo de regresión, sería muy laborioso e desarrollo de nuevos modelos sobre todo porque las redes neuronales tardan mucho en aprender y los modelos de regresión no capturarían las posibles correlaciones no lineales que puede existir entre los registros. Con esta red neuronal se puede predecir el suceso de eventos ponderándolos por las veces en que se han presentado en los datos históricos, por un lado es positivo porque puede especializarse en una ventana de tiempo específico, por otro lado esta especialización puede hacer que olvide ciertos eventos que son muy esporádicos y solamente preste atención a eventos muy específicos. En futuras versiones se incluirán algunas condiciones ambientales como la temperatura y algunas condiciones socioeconómicas de los cuadrantes geográficos.

Bibliografía

- [1] Tayal, Devendra Kumar, Jain, Arti, Arora, Surbhi, Agarwal, Surbhi, Gupta, Tushar Tyagi, Nikhil. Crime detection and criminal identification in India using data mining techniques. AI & SOCIETY. Vol. 30, pp. 117-127. 2015
- [2] Gerber, Matthew S. Predicting crime using Twitter and kernel density estimation. Decision Support Systems. Vol. 61, pp. 115-125. 2014
- [3] Ku, Chih-Hao, Leroy, Gondy. A decision support system: Automated crime report analysis and classification for e-government. Government Information Quarterly. Vol. 31, Issue 4, pp. 534-544. 2014
- [4] J. Park. A Study of the Probability of Prediction to Crime according to Time Status Change. Journal of The Korea Society of Computer and Information. Vol. 18, Issue 5, pp. 147-156. 2013
- [5] Oatley, Giles, Ewart, Brian. Data mining and crime analysis. Wiley Interdisciplinary Reviews-Data Mining And Knowledge Discovery. Vol. 1, issue 2, pp. 147-153. 2011

- [6] Johnson, Shane D. A brief history of the analysis of crime concentration
European Journal Of Applied Mathematics. Vol. 21, Issue 4-5, pp. 349-370
2010
- [7] Helbich, Marco, Hagenauer, Julian, Leitner, Michael, Edwards, Ricky. Exploration of unstructured narrative crime reports: an unsupervised neural network and point pattern analysis approach. *Cartography And Geographic Information Science*. Vol 40, Issue 4, pp. 326-336. (2013)
- [8] Li, Sheng-Tun, Kuo, Shu-Ching, Tsai, Fu-Ching. An intelligent decision-support model using FSOM and rule extraction for crime prevention. *Expert Systems With Applications*. Vol. 37, Issue 10, pp. 7108-7119. 2010
- [9] Kianrneh, Keivan; Alhajj, Reda. Effectiveness of support vector machine for crime hot-spots prediction. *Applied Artificial Intelligence*. Vol. 22, Issue 5, pp. 433-458. 2008
- [10] Alwee, Razana, Shamsuddin, Siti Mariyam Hj, Sallehuddin, Roselina
Hybrid Support Vector Regression and Autoregressive Integrated Moving Average Models Improved by Particle Swarm Optimization for Property Crime Rates Forecasting with Economic Indicators. *Scientific World Journal*
Article number 951475. 2013
- [11] Sanmiguel, G.G.; Gonzalez, L.L.; Torres-Trevino, L.M.; Guerra, C., "On-Line Learning in an Embedded Maximum Sensibility Neural Network," in *Artificial Intelligence (MICAI), 2012 11th Mexican International Conference on*, vol., no., pp.75-79, Oct. 27 2012-Nov. 4 2012
doi: 10.1109/MICAI.2012.19
- [12] Escamilla, Indira and Torres, Luis and Perez, Pedro and Zambrano, Patricia
A Comparison between Back Propagation and the Maximum Sensibility Neural Network to Surface Roughness Prediction in Machining of Titanium (Ti 6Al 4V) Alloy
Lecture Notes in Computer Science, Springer
Vol. 5317, pp. 1009-1019, 2008

Acerca de los Autores

Francisco Torres
Leticia Neira
Luis Torres-Treviño
FIME Posgrado – UANL.
San Nicolás de los Garza, N.L. México