

Aplicación de escritorio de apoyo para la predicción de asaltos a transeúntes y gestión de datos derivados en la delegación Gustavo A. Madero

Trabajo terminal No. 2021-A097

Alumnos: Martínez Gallegos Jesús
Directores: Cruz Meza María Elena
e-mail: jmartinezg1208@alumno.ipn.mx

Resumen – Durante 2019, a pesar de presentarse una reducción en la incidencia de robos a transeúnte, estos siniestros siguen siendo el delito más común en la Ciudad de México (CDMX), representando un 38.6% de todos los registrados [1]. Por esta razón, se pretende desarrollar una aplicación de escritorio que implemente técnicas de inteligencia artificial para predecir la incidencia del delito que, al ser utilizada por instituciones de seguridad pública, apoye en la administración de los recursos de vigilancia disponibles en campo, indicando los lugares o zonas más propensas a presentar incidentes de robo a transeúnte.

Palabras clave – Aplicación de escritorio, Inteligencia artificial, Predicción, Robo a transeúnte

1. Introducción

En 2020, el Observatorio Nacional Ciudadano (ONC) informó que la Ciudad de México ocupó el primer lugar en robo a transeúntes a nivel Nacional [2] siendo que este delito es reportado por 23,911 de cada 100,000 habitantes [1]. El resultado de estos crímenes afecta directamente a la víctima, así como al resto de la población generando una noción de inseguridad.

Fundamentalmente, la prevención del crimen requiere de un incremento en la vigilancia, esto implica invertir en recursos tanto económicos como humanos que no siempre es posible solventar. La predicción de delitos ha recibido una considerable atención dados los beneficios que promete. Aplicar los resultados de la predicción a la administración de recursos generaría patrullas más eficaces. Este es un tema relevante en la CDMX dado que las instituciones de seguridad pública cuentan con una cantidad limitada de personal que debe de ser distribuido estratégicamente para obtener la mayor eficacia posible.

Actualmente existen sistemas que son capaces de predecir la incidencia de delitos, pero su implementación está limitada a zonas específicas en países primermundistas, además, su costo es muy elevado o en otros casos simplemente no es posible obtener acceso a ellos. Un ejemplo de estos sistemas es PredPol, que es un software de escritorio para la “predicción policial”, este analiza datos sobre la incidencia de delitos a través de un algoritmo de machine learning el cual registra los datos históricos de los crímenes de cada zona para determinar los lugares más probables en donde podría ocurrir el próximo acto delictivo, el costo de este software por año se encuentra entre los 10,000 y 150,000 dólares por año.[4]

En México existen sistemas como el Comando, Control, Cómputo, Comunicaciones y Contacto Ciudadano de la Ciudad de México (C5) para la vigilancia y administración del personal que desde el año 2009 permite mejorar la reacción de las autoridades ante emergencias, situaciones de crisis y comisión de ilícitos. [5] Este tipo de sistemas no cuentan con la generación de predicciones, las cuales han ayudado a reducir los índices delictivos hasta en un 30% en ciudades de los Estados Unidos [6].

Sistemas similares que se han desarrollado son:

1. PredPol
2. CompStat
3. hoyodecrimen.com
4. TT2020-A108

Tabla 1: Resumen de productos similares

Proyecto	Características	Precio en el mercado
PredPol	Usa historial estadístico. Enfocado a uso policial. Aplicación de escritorio. Genera predicciones.	\$10,000 - \$150,000 dólares por año
CompStat	Usa historial estadístico. Enfocado a uso policial. Sistema parecido a una base de datos. Genera predicciones.	Uso reservado a los departamentos de policía de Estados Unidos
hoyodecrimen.com	Página web Muestra mapas con los puntos de incidencia Reúne información estadística de delitos en la ciudad de México dividida por cuadrantes Cuenta con una API pública	Gratuito
TT2020-A108 [7]	Análisis de datos Modelo de predicciones. Trabajo terminal de ESCOM.	

Es importante hacer notar que con este trabajo se intenta cubrir uno de los puntos tratados como trabajo futuro en el trabajo terminal TT2020-A108 puesto que lo ideal sería contar con una herramienta completa, al proponer una aplicación para escritorio se pretende implementar la estrategia propuesta en el TT 2020-A108 [7] ya que en dicho trabajo se define los datos de las fuentes de interés y las técnicas para la predicción, sin embargo, hay otras técnicas en el ámbito de Machine Learning que vale la pena estudiar en conjunto con el análisis predictivo y el análisis de sentimientos para extraer información crítica de los conjuntos de datos recopilados.

Con lo anterior, se ve la necesidad de desarrollar una aplicación que permita la gestión de los datos que el trabajo terminal 2020-A108 ha logrado concentrar. La gestión de los datos en este tipo de problemas es de suma importancia a nivel directivo ya que la analítica de datos permite una mejor toma de decisiones, un ejemplo puede ser que la herramienta permita apoyar a los directivos de la alcaldía GAM a presentar una mejor distribución del cuerpo policiaco en dicha zona.

2. Objetivo

Desarrollar una aplicación de escritorio que funcione como una herramienta de apoyo para gestionar consultas detalladas de las zonas con una alta probabilidad de ocurrencia de robos a transeúntes en la alcaldía Gustavo A. Madero (GAM) de la Ciudad de México (CDMX)(análisis de los datos del periodo 2016-2021).

Objetivos específicos:

- Recolección de requerimientos
- Definición de herramientas de desarrollo
- Análisis y recolección de datos de los crímenes ocurridos en el periodo 2016-2021
- Análisis y diseño del prototipo
 - Investigación de las técnicas de Inteligencia Artificial y Ciencia de Datos
- Diseño e implementación de las Técnicas de Inteligencia Artificial y Ciencia de Datos
- Implementación de la capa de presentación de datos

- Pruebas e integración de los módulos del prototipo

3. Justificación

En la actualidad existen diversos sistemas que pretenden predecir distintos actos delictivos, desafortunadamente estos solo han sido desarrollados en países potencia dado su costo, como es el caso de PredPol que tiene un precio que va desde los \$10,000 hasta los \$150,000 dólares. Estas herramientas se basan en suposiciones sobre el riesgo de delincuencia y sus relaciones con el tiempo, el entorno y la sociedad. La capacidad de predecir una futura ubicación en los puntos críticos de delincuencia trae innumerables ventajas a una fuerza policial, incluyendo la planificación de intervenciones proactivas efectivas, reducir el nivel de delincuencia, promover la seguridad pública y mejorar la eficiencia de los recursos policíacos [6].

La implementación de un sistema de este tipo en la CDMX ayudaría a organizar de la mejor manera los limitados elementos de seguridad que se tienen en campo y así obtener resultados prometedores a favor de la seguridad pública de los transeúntes.

Se eligió la alcaldía Gustavo A. Madero para la implementación de este proyecto ya que es una de la más inseguras de la CDMX con altos índices delictivos que van en incremento anualmente.[3] En esta alcaldía se realizan diferentes actividades para la prevención de delito como son conferencias a sus ciudadanos, operativos para la revisión de transeúntes para evitar la portación de armas, monitoreo por cámaras de vigilancia, botones de pánico, entre algunas otras [8]; lo que indica una alta iniciativa en la mejora de la seguridad ciudadana.

Para ayudar a solventar este problema, se propone desarrollar una aplicación de escritorio que, usando análisis de datos y técnicas de inteligencia artificial, prediga las ubicaciones más probables a presentar un asalto a transeúnte mediante la recopilación de datos estadísticos de los crímenes ocurridos en los últimos 5 años. Además, proponemos que esta información sea plasmada en mapas interactivos, similar a un mapa de calor, permitiendo que estos datos sean más fáciles de interpretar. Con esta información como herramienta se pretende gestionar los recursos de patrullaje de la Alcaldía GAM con mayor efectividad y con ello reducir los robos a transeúnte considerablemente.

En la actualidad, existe una inmensa cantidad de información que es recopilada día a día y en la última década, gracias al rápido desarrollo de hardware de cómputo, se ha popularizado el uso del análisis de datos para aprovechar estos volúmenes de información tan grandes. El análisis de datos proporciona métodos para agilizar el manejo de grandes conjuntos de datos basándose en el área de la estadística, permitiendo obtener la información más relevante del conjunto. [9] En compañía de las técnicas de inteligencia artificial se ha logrado generar modelos muy flexibles que pueden ser aplicados en múltiples ámbitos simplemente cambiando el conjunto de datos de entrenamiento.

La ESCOM cuenta con unidades de aprendizaje que introducen los principios de inteligencia artificial, impartidos por docentes con experiencia en estas áreas. Estos recursos son de mucha ayuda al proponer un proyecto de este tipo ya que otorgan el entendimiento básico para analizar e implementar modelos más complejos además de contar con el apoyo de personal conocedor del área.

A continuación se muestra una propuesta de la arquitectura de la aplicación:

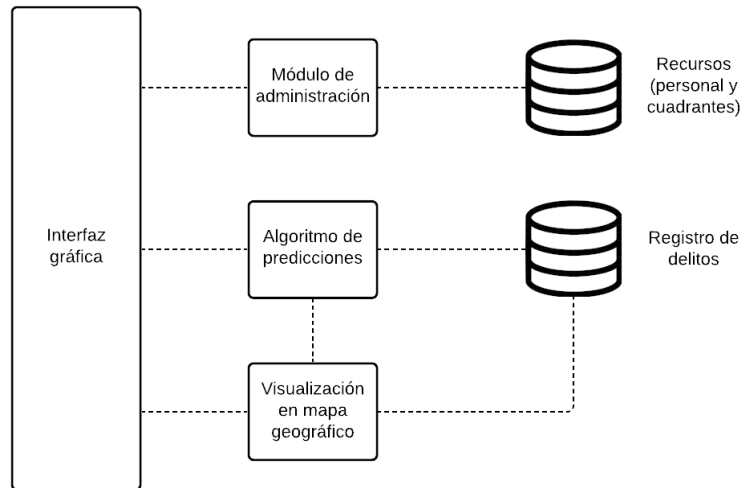


Ilustración 1: Arquitectura propuesta (a grandes rasgos)

La aplicación se basará en tres módulos principales:

- El módulo de administración permitirá acceder al listado de personal disponible en campo y se les podrá asignar manualmente a un cuadrante de la delegación.
- El módulo de predicciones, que contará con dos fases de desarrollo:
 - Entrenamiento, en donde se hará uso de los datos históricos de delitos de robo a transeúnte registrados (2016-2021) para generar un modelo predictivo.
 - Generación de predicciones, será el modelo ya entrenado y de acuerdo con la información de una zona en la delegación generará una predicción de la probabilidad de ocurrencia de un delito de robo a transeúnte.
- El módulo del mapa geográfico mostrará una especie de mapa de calor en el cual se indicará las zonas más probables a presentar un robo a transeúnte, esto con la ayuda del módulo de predicción. Además, se podrá visualizar información histórica como la ubicación de delitos anteriores.

Desarrollar una aplicación con estas características no es un trabajo fácil ya que es necesario hacer uso de múltiples áreas de aprendizaje para llevarla a cabo. Sin duda, las tareas más demandantes serán la selección de los datos relevantes, la investigación de los modelos más apropiados para generar predicciones acertadas y su correcta implementación. Además, la etapa de entrenamiento de dichos modelos requerirá de un gran poder de cómputo, para ello existen alternativas de pago, pero de momento, se considera el uso de Google Colaboratory que es una herramienta gratuita que cubriría las necesidades para esta fase del proyecto.

4. Productos o Resultados esperados

Al finalizar el trabajo terminal, los entregables serán:

- Código fuente
- Aplicación de escritorio
- Manual de usuario
- Manual técnico

5. Metodología

Se utilizará el marco de trabajo por prototipos para este proyecto, específicamente su variante incremental. Siguiendo este modelo, se construye un prototipo dividido en distintos módulos Ilustración 2, una vez diseñado, se puede trabajar paralelamente en cada uno de estos módulos, sin embargo, en este proyecto se realizará secuencialmente dada la cantidad de integrantes en el equipo.

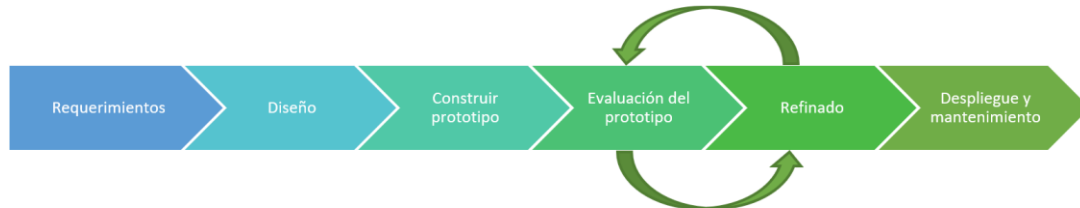


Ilustración 2 Etapas del modelo por prototipos

Un modelo de creación de prototipos comienza con el análisis de requisitos. En esta fase se definen en detalle los requisitos del sistema. Durante el proceso se entrevista a los usuarios del sistema para saber cuál es su expectativa del sistema. La segunda fase es un diseño preliminar o un diseño rápido donde se crea un diseño simple del sistema, sin embargo, no es un diseño completo, pero da una breve idea del sistema al usuario.

En este momento se entra a un ciclo de evaluación y refinado; si el usuario no está satisfecho con el prototipo actual, se debe refinar el prototipo de acuerdo con los comentarios y sugerencias. Esta fase no terminará hasta que se cumplan todos los requisitos especificados por el usuario. Una vez que el usuario está satisfecho con el prototipo desarrollado, se desarrolla un sistema final basado en el prototipo final aprobado.

Una vez que el sistema final se desarrolla en base al prototipo final, se prueba exhaustivamente y se implementa en producción. El sistema se somete a un mantenimiento de rutina para minimizar el tiempo de inactividad y evitar fallas a gran escala [10].

Las herramientas que se pretenden utilizar son:

- Back end
 - Pytorch
 - Tensorflow
 - Perceptilabs
 - Google Colaboratory
- Front end
 - WxWidgets
 - PyQt5
 - Google Maps API
- Otras herramientas para la comunicación y gestión de documentos son:
 - Para la comunicación entre los miembros involucrados (alumno, directora del TT, sinodales y docente de seguimiento que se asigne) en el desarrollo del TT se pretende utilizar Microsoft Teams.
 - Para la gestión del proyecto se usará tara.ai
 - Para la documentación del trabajo se usará app.diagrams.net

6. Cronograma

Actividad	AGO	SEP	OCT	NOV	DIC	ENE	FEB	MAR	ABR	MAY	JUN
Mapeado de requerimientos funcionales											
Mapeado de requerimientos no funcionales											
Bosquejo de casos de uso											
Definición de las características del modelo de inteligencia artificial											
Recolección de datos sin formato											
Modelado de datos de entrenamiento											
Diseño de las bases de datos											
Diseño de interfaces gráficas											
Generación de diagramas de la arquitectura											
Selección del modelo de IA											
Selección de herramientas según la arquitectura propuesta											
Crear las bases de datos											
Escribir el código para el modelo de IA											
Escribir el código para el módulo administrativo											
Escribir el código para la interfaz gráfica											
Etapas de pruebas											
Generación de manuales											

7. Referencias

- [1] INEGI. (2020). Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2020. [Online]. INEGI Sitio web: https://www.inegi.org.mx/contenidos/programas/envipe/2020/doc/envipe2020_cdmx.pdf
- [2] C. Lara, "CDMX, primer lugar en robo a transeúnte a nivel nacional", El Sol De Mexico, 2020. [Online]. Sitio web: <https://www.elsoldemexico.com.mx/metropoli/policia/cdmx-primero-lugar-en-robo-a-transeunte-a-nivel-nacional-disminucion-delitos-observatorio-nacional-ciudadano-5164557.html>
- [3] onseguridad. (2018). Las 5 alcaldías más y menos seguras de la CDMX durante 2018. 2020, de inseguridad Sitio web: <https://lopezdoriga.com/wp-content/uploads/2019/06/recomendaciones.pdf>
- [4] "PredPol Mission | About Us | Aiming to reduce victimization keep communities safer", PredPol, 2020. [Online]. Sitio web: <https://www.predpol.com/about/>
- [5] C5 de la CDMX, 2018. [Online]. Available: <https://www.c5.cdmx.gob.mx/dependencia/acerca-de/el-c5-de-la-cdmx>
- [6] Ellen Huet. (abril 8,2015). PredPol, el primer software de predicción del crimen. 2020. [Online]. FORBES Sitio web: <https://www.forbes.com.mx/predpol-el-primero-software-de-prediccion-del-crimen/>
- [7] O. Montes, "Prototipo de Herramienta para la predicción de asaltos a transeúntes en la alcaldía Gustavo A. Madero de la Ciudad de México," 2020. TT2020-A108

- [8] “Dirección Ejecutiva de Seguridad Ciudadana, Gestión Integral de Riesgos y Protección Civil,” Alcaldía Gustavo A. Madero, 2019. [Online]. Sitio web: <http://www.gamadero.gob.mx/SeguridadCiudadana/>
- [9] M. Miyazaki, “A Brief History of Data Analysis,” FlyData, 2020. [Online]. Sitio web: <https://www.flydata.com/blog/a-brief-history-of-data-analysis/>
- [10] M. Martin, «Prototyping Model in Software Engineering: Methodology, Process, Approach,» [En línea]. Available: <https://www.guru99.com/software-engineering-prototyping-model.html>

8. Alumnos y directores

Martínez Gallegos Jesús. - Alumno de la carrera de Ing. en Sistema Computacionales en ESCOM, Especialidad Sistemas, Boleta 2013070818. Tel. 5533175130, email jmartinezg1208@alumno.ipn.mx
Firma: _____

Cruz Meza María Elena. M. en C. de la Computación, en la especialidad de Reconocimiento de Patrones y Tratamiento Digital de Imágenes por el Centro de Investigación en Computación CIC del IPN. Áreas de interés: aprendizaje de máquina, análisis de imágenes, visión por computadora, ciencia de datos, algoritmos genéticos. Tel.57296000 ext. 52032, email: mcruz@ipn.mx
Firma: _____

CARÁCTER: Confidencial
FUNDAMENTO LEGAL: Artículo 11 Fracc. V y Artículos 108, 113 y 117 de la Ley Federal de Transparencia y Acceso a la Información Pública.
PARTES CONFIDENCIALES: Número de boleta y teléfono.

← Responder ↶ Responder a todos → Reenviar ⚙ 📁 Archivar 🗑 Eliminar 📌 Establecer marca ⋮

RE: Protocolo de Investigación: Predicción de asalto a transeúntes



Maria Elena Cruz Meza <mcruzm@ipn.mx>

07:08 p. m.



Para: Jesus Martinez Gallegos

Acuso de recibido.

Saludos cordiales Jesús!

María Elena Cruz Meza
Profesora
Academia de Ing. de SW
Depto. de ISC.
Escom.IPN

De: Jesus Martinez Gallegos <jmartinezg1208@alumno.ipn.mx>

Enviado: viernes, 4 de junio de 2021 20:51

Para: Maria Elena Cruz Meza <mcruzm@ipn.mx>

Asunto: RE: Protocolo de Investigación: Predicción de asalto a transeúntes