

VICTIMAS EN CARPETAS DE INVESTIGACIÓN DE LA FISCALÍA GENERAL DE JUSTICIA DE LA CIUDAD DE MÉXICO

Flor Margarita Calderón Cuevas

Escuela Superior De Ciencias y Tecnologías De La Información

Correo Electrónico: florcalderon@uagro.mx

RESUMEN

Este artículo tiene como propósito presentar un trabajo de minería de datos el cual se realizó gracias a la descarga de una base de datos del Portal de datos de la ciudad de México; la cual contiene las carpetas de investigación que se abrieron en los años 2019 y 2020, en ellas se tienen los registros de las fechas, horas, la alcaldía y la colonia de los hechos; además el tipo de delito, el sexo y la edad de la víctima, así como el tipo de persona ya sea física o moral, etc.

Palabras clave: *minería de datos, México, Datos, Weka, RStudio*

ABSTRACT

The purpose of this article is to present a data mining work which was carried out thanks to a database that was downloaded from the Mexico City Data Portal; The database contains the investigation folders that were opened in the years 2019 and 2020, in them there are the records of and the colony of the events; also the type of crime, sex and age of the victim, as well as the type of person, whether physical or moral, etc. the dates, times, the mayor's office, etc.

Keywords: *Data mining, Mexico, Data, Weka, RStudio*

1. INTRODUCCION

La ciudad de México está ubicada en la zona central de la República Mexicana, en la parte sur del valle de México, tiene una extensión de 1.495 kilómetros cuadrados (Km²) y ocupa el lugar 32 a nivel nacional; es decir que es la entidad más pequeña de la República Mexicana; la Ciudad de México representa 0.08% de la superficie del país, colinda al norte, al este y al oeste con el estado de México y al sur con el estado de Morelos. De acuerdo con el INEGI sobre la tasa de incidencia delictiva por entidad federativa de ocurrencia por cada 100 mil habitantes, la ciudad de México en los años 2017 y 2018 registro una tasa del 68,954 y 69,716, es decir mayor, la cual se llevó el primer lugar en la entidad federativa con el mayor número de delitos reportados (INEGI, 2018). La ciudad de México ofrece un portal de datos en cual se pueden descargar y explorar bases de datos; dentro del portal se obtuvo la de Víctimas en carpetas de investigación FGJ, con la cual se le aplico un trabajo de minería de datos.

1.1 OBJETIVO GENERAL

El objetivo del trabajo es que, con la base de datos obtenida del portal de datos de la ciudad de México, se pueda generar información estadística para obtener una mejor visión de la situación actual de la delincuencia en la que se encuentra la ciudad; además de ofrecer información preventiva acerca de las alcaldías y colonias con un número mayor de delitos, para que los ciudadanos tomen conciencia de ello.

2. METODOLOGÍA

Para llevar a cabo el presente trabajo se utilizó una serie de fases las cuales son: integración y recopilación, selección, limpieza y transformación, minería de datos, pruebas y verificación de resultados. (Valero Orea, Salvador Vargas, & Garcia Alonso, 2005)

2.1 Integración y recopilación

La base de datos fue obtenida del Portal de datos de la ciudad de México (México, Portal de datos de la Ciudad de México), se seleccionó el tema de Víctimas en carpetas de investigación FGJ, se obtuvo todo el conjunto de datos en formato .CSV además del diccionario de los datos. Contiene registro de las víctimas de los delitos en las carpetas de la Fiscalía de General de Justicia (FGJ) de la ciudad de México a partir de enero de 2019 y 2020.

Contenía los siguientes campos:

- **idCarpeta:** Es un número entero que representa el identificador único usado por FGJ asociado a cada carpeta de investigación dentro de su sistema.
- **Año_Inicio:** Año en que se hizo la denuncia para iniciar la carpeta de investigación.
- **Mes_Inicio:** Mes en que se hizo la denuncia para iniciar la carpeta de investigación.
- **Fecha_Inicio:** Día en que se hizo la denuncia para iniciar la carpeta de investigación.
- **Hora_Inicio:** Hora y minuto en que se hizo la denuncia para iniciar la carpeta de investigación.
- **Delito:** Es la conducta, acción u omisión típica (descrita por la ley), antijurídica (contraria a la ley) y culpable, a la que le corresponde una sanción.
- **Categoría:** Las carpetas de investigación se clasifican en función del tipo de delito cometido, los cuales se dividen en 16.
- **Sexo:** Sexo de la víctima del delito reportado.
- **Edad:** Edad de la víctima del delito reportado en la carpeta de investigación.
- **TipoPersona:** Cómo se reconoce a él o los denunciados al ser sujetos de obligaciones y derechos salvaguardados.
- **CalidadJuridica:** Título con el que se registra y se identifica a la persona, física o moral, en la carpeta de investigación.
- **Competencia:** Variable categórica a través de la cual se clasifican los hechos según su naturaleza.

- Hechos no delictivos corresponde a aquellos que son denunciados a la PGJ, pero no constituyen un delito en sí mismos como por ejemplo un suicidio.
 - Incompetencias son aquellos hechos delictivos que suceden fuera de la Ciudad de México y son denunciados a la PGJ de la Ciudad de México, por lo que no se deben tomar en cuenta como incidencia delictiva propia de la ciudad.
 - Fuero común son los delitos que ocurren y se denuncian dentro de la Ciudad de México.
 - Año_Hecho: Año en que se cometió el delito, según el reporte de la víctima.
 - Mes_Hecho: Mes en que se cometió el delito, según el reporte de la víctima.
 - Fecha_Hecho: Día en que se cometió el delito, según el reporte de la víctima.
 - Hora_Hecho: Hora y minuto en que se cometió el delito, según el reporte de la víctima.
 - AlcaldíaHechos: Alcaldía en que se cometió el delito, según el reporte de la víctima. Notar que puede ser fuera de la CDMX.
 - ColoniaHechos: Colonia en que se cometió el delito, según el reporte de la víctima.
 - Calle_Hechos: Calle en que se cometió el delito, según el reporte de la víctima.
 - Calle_Hechos2: Calle secundaria en que se cometió el delito, según el reporte de la víctima.
 - Lon: Longitud de la geolocalización, uno de dos elementos que componen la referencia angular que permite localizar el lugar donde se cometió el delito. WGS84.
 - Lat: Latitud de la geolocalización, uno de dos elementos que componen la referencia angular que permite localizar el lugar donde se cometió el delito. WGS84.
- Cita: (México, Portal de datos de la ciudad de México)

2.2 Selección, limpieza y transformación

Después de haber llevado a cabo la investigación y análisis de los campos que se encontraban en la base de datos (ya descritos anteriormente) se prosiguió a seleccionar los campos útiles para la investigación; para ello la base de datos en formato .CSV fue convertida al formato Excel, ya que con ese formato fue más fácil la selección y limpieza de los datos; Excel es una herramienta fácil de manejar en cuestiones de búsqueda y reemplazo de algunos datos, ya que algunos campos que contenía la base se encontraban con errores ortográficos ocasionando problemas en Weka, ya que Weka fue una plataforma de software seleccionada para trabajar la parte de minería de datos al igual que RStudio como se puede observar en la figura 1 y 2.

La base de datos fue limpiada y organizada de tal forma que se permitiera hacer uso de ella en las plataformas de Weka y RStudio. Al haber transformado la base de datos en formato Excel y hacer todo el procedimiento correspondiente a la selección, limpieza y transformación de los datos, se volvió a convertir en formato .CSV ya que con ese tipo de formato es más fácil trabajar con las plataformas mencionadas anteriormente.

Viewer

Relation: datosnuevos1-weka.filters.supervised.attribute.Discretize-Rfirst-last-precision6-weka.filters.supervised.attribute.Discretize-Rfirst-last-precision6

No.	1: i>Anio_inicio Nominal	2: Mes_inicio Nominal	3: Delito Nominal	4: Categoria Nominal	5: Sexo Nominal	6: Edad Nominal	7: Anio_Hecho Nominal	8: Mes_Hecho Nominal	9: Hora_Hecho Nominal	10: Alcaldia_Hechos Nominal	11: Colonia_Hechos Nominal
1	'All'	Abril	NARCO...	DELITO DE ...	NA	NA	'All'	Marzo	23:20:00	GUSTAVO A MAD...	LOMA LA PALMA
2	'All'	Abril	NARCO...	DELITO DE ...	NA	NA	'All'	Marzo	23:20:00	GUSTAVO A MAD...	LOMA LA PALMA
3	'All'	Abril	ROBO ...	ROBO A TR...	Mascul...	34	'All'	Marzo	19:40:00	IZTAPALAPA	VALLE DEL SUR
4	'All'	Abril	FRAUDE	DELITO DE ...	NA	NA	'All'	Enero	12:00:00	BENITO JUAREZ	NAPOLES
5	'All'	Abril	FRAUDE	DELITO DE ...	Mascul...	NA	'All'	Enero	12:00:00	BENITO JUAREZ	NAPOLES
6	'All'	Abril	PRIVA...	DELITO DE ...	Mascul...	NA	'All'	Marzo	11:45:00	ALVARO OBREGON	LOMAS DE SANTA FE
7	'All'	Abril	DAÑO ...	DELITO DE ...	Mascul...	22	'All'	Marzo	22:20:00	GUSTAVO A MAD...	PALMATITLA
8	'All'	Abril	TENTA...	DELITO DE ...	Mascul...	26	'All'	Marzo	17:00:00	XOCHIMILCO	PISTA OLIMPICA VI...
9	'All'	Abril	ROBO ...	ROBO A NE...	Femen...	34	'All'	Marzo	20:00:00	MIGUEL HIDALGO	VERONICA ANZURES
10	'All'	Abril	ROBO ...	ROBO A NE...	NA	NA	'All'	Marzo	20:00:00	MIGUEL HIDALGO	VERONICA ANZURES
11	'All'	Abril	CORRU...	DELITO DE ...	Mascul...	6	'All'	Marzo	12:00:00	GUSTAVO A MAD...	NUUEVA VALLEJO
12	'All'	Abril	AMENA...	DELITO DE ...	Femen...	47	'All'	Marzo	19:00:00	TLALPAN	LOMAS DE PADERNA
13	'All'	Abril	ROBO ...	DELITO DE ...	NA	NA	'All'	Marzo	15:00:00	CUAJIMALPA DE ...	BOSQUES DE LAS L...
14	'All'	Abril	ROBO ...	DELITO DE ...	NA	NA	'All'	Marzo	15:00:00	CUAJIMALPA DE ...	BOSQUES DE LAS L...
15	'All'	Abril	VIOLA...	VIOLACION	Femen...	26	'All'	Marzo	0:00:00	TLALPAN	ISIDRO FABELA
16	'All'	Abril	VIOLA...	DELITO DE ...	Mascul...	25	'All'	Marzo	23:30:00	AZCAPOTZALCO	SANTIAGO AHUIZO...
17	'All'	Abril	VIOLA...	DELITO DE ...	Femen...	36	'All'	Marzo	23:30:00	AZCAPOTZALCO	SANTIAGO AHUIZO...
18	'All'	Abril	FRAUDE	DELITO DE ...	Femen...	NA	'All'	Febrero	12:00:00	BENITO JUAREZ	DEL VALLE CENTRO
19	'All'	Abril	FRAUDE	DELITO DE ...	NA	NA	'All'	Febrero	12:00:00	BENITO JUAREZ	DEL VALLE CENTRO
20	'All'	Abril	ROBO ...	DELITO DE ...	Mascul...	31	'All'	Marzo	16:00:00	MIGUEL HIDALGO	ESCONDON I SECCI...
21	'All'	Abril	VIOLA...	DELITO DE ...	Femen...	26	'All'	Marzo	23:00:00	MILPA ALTA	SAN SALVADOR CU...
22	'All'	Abril	FRAUDE	DELITO DE ...	Mascul...	NA	'All'	Febrero	15:00:00	MIGUEL HIDALGO	POLANCO
23	'All'	Abril	FRAUDE	DELITO DE ...	Femen...	NA	'All'	Enero	12:00:00	BENITO JUAREZ	SAN PEDRO DE LOS...
24	'All'	Abril	VIOLA...	DELITO DE ...	Mascul...	46	'All'	Marzo	23:00:00	VENUSTIANO CA...	LORENZO BOTURINI
25	'All'	Abril	VIOLA...	DELITO DE ...	Femen...	37	'All'	Marzo	19:00:00	BENITO JUAREZ	PORTALES SUR
26	'All'	Abril	DESPOJO	DELITO DE ...	Mascul...	74	'All'	Marzo	23:40:00	ALVARO OBREGON	PUEBLO SANTA FE
27	'All'	Abril	FRAUDE	DELITO DE ...	Femen...	NA	'All'	Marzo	12:00:00	BENITO JUAREZ	NAPOLES
28	'All'	Abril	FRAUDE	DELITO DE ...	NA	NA	'All'	Marzo	12:00:00	BENITO JUAREZ	NAPOLES
29	'All'	Abril	ABUSO...	DELITO DE ...	Femen...	38	'All'	Marzo	12:30:00	VENUSTIANO CA...	MOCTEZUMA 2A SE...
30	'All'	Abril	DENUN...	HECHO NO ...	Femen...	21	'All'	Marzo	1:00:00	AZCAPOTZALCO	CENTRO DE AZCAP...

Add instance Undo OK Cancel

Ilustración 1 Base de datos en Weka

DELITOS.R x delinuencia x

Filter

	LidCarpeta	Anio_inicio	Mes_inicio	FechaInicio	Delito	Categoria	Sexo	Edad	TipoPersona
1	8393499	2019	Abril	01/04/2019	NARCOMENUEDO POSESION SIMPLE	DELITO DE BAJO IMPACTO	NA	NA	NA
2	8393499	2019	Abril	01/04/2019	NARCOMENUEDO POSESION SIMPLE	DELITO DE BAJO IMPACTO	NA	NA	MORAL
3	8393500	2019	Abril	01/04/2019	ROBO A TRANSEUNTE EN VIA PUBLICA CON VIOLENCIA	ROBO A TRANSEUNTE EN VIA PUBLICA CON Y SIN VIOLEN...	Masculino	34	FISICA
4	8393502	2019	Abril	01/04/2019	FRAUDE	DELITO DE BAJO IMPACTO	NA	NA	MORAL
5	8393502	2019	Abril	01/04/2019	FRAUDE	DELITO DE BAJO IMPACTO	Masculino	NA	FISICA
6	8393503	2019	Abril	01/04/2019	PRIVACION DE LA LIBERTAD PERSONAL	DELITO DE BAJO IMPACTO	Masculino	NA	FISICA
7	8393504	2019	Abril	01/04/2019	DAÑO EN PROPIEDAD AJENA CULPOSA POR TRANSITO VE...	DELITO DE BAJO IMPACTO	Masculino	22	FISICA
8	8393505	2019	Abril	01/04/2019	TENTATIVA DE EXTORSION	DELITO DE BAJO IMPACTO	Masculino	26	FISICA
9	8393506	2019	Abril	01/04/2019	ROBO A NEGOCIO CON VIOLENCIA	ROBO A NEGOCIO CON VIOLENCIA	Femenino	34	FISICA
10	8393506	2019	Abril	01/04/2019	ROBO A NEGOCIO CON VIOLENCIA	ROBO A NEGOCIO CON VIOLENCIA	NA	NA	MORAL
11	8393507	2019	Abril	01/04/2019	CORRUPCION DE PERSONAS MENORES DE EDAD O PERSO...	DELITO DE BAJO IMPACTO	Masculino	6	FISICA
12	8393508	2019	Abril	01/04/2019	AMENAZAS	DELITO DE BAJO IMPACTO	Femenino	47	FISICA
13	8393509	2019	Abril	01/04/2019	ROBO A NEGOCIO SIN VIOLENCIA	DELITO DE BAJO IMPACTO	NA	NA	MORAL
14	8393509	2019	Abril	01/04/2019	ROBO A NEGOCIO SIN VIOLENCIA	DELITO DE BAJO IMPACTO	NA	NA	MORAL
15	8393511	2019	Abril	01/04/2019	VIOLACION	VIOLACION	Femenino	26	FISICA
16	8393512	2019	Abril	01/04/2019	VIOLENCIA FAMILIAR	DELITO DE BAJO IMPACTO	Masculino	25	FISICA
17	8393512	2019	Abril	01/04/2019	VIOLENCIA FAMILIAR	DELITO DE BAJO IMPACTO	Femenino	36	FISICA
18	8393513	2019	Abril	01/04/2019	FRAUDE	DELITO DE BAJO IMPACTO	Femenino	NA	FISICA

Showing 1 to 20 of 429,678 entries, 18 total columns

Ilustración 2 Base de datos en RStudio

2.3 Minería de datos

En esta fase se realizó un documento en Word en el cual se especificaba que era lo que se tenía planeado hacer con la base de datos, es decir que algoritmos o métodos se iban a utilizar; así como las plataformas de software con las que se iba trabajar. Se realizó la selección de las plataformas dependiendo de las actividades que estaban planteadas en el documento, es decir ver donde convendría hacer cada actividad si en Weka o en RStudio.

Weka es una plataforma de software de libre distribución desarrollado en Java. Está constituido por una serie de paquetes de código abierto con diferentes técnicas de preprocesado, clasificación, agrupamiento, asociación, y visualización, así como facilidades para su aplicación y análisis de prestaciones cuando son aplicadas a los datos de entrada seleccionados (Corso, 2009).

Por lo anterior mencionado y por otras más razones fue que Weka fue seleccionada para trabajar con la base de datos de esta investigación. En RStudio fue diferente, ya que seleccionada esta plataforma debido a la poca experiencia que se tenía, pero la cual era la necesaria para realizar esta investigación.

Las actividades que se plantearon en el documento en Word fueron las siguientes:

- Comparar los años en este caso 2019 y 2020
- Relación entre los meses
- Comparación entre los delitos
- La categoría de los delitos
- El sexo y la edad
- Que tipo de persona moral y física
- La calidad jurídica
- Hora de los hechos
- Las colonias y la relación con los delitos
- Reglas de asociación
- Generación de gráficos entre la hora y el mes
- Utilizar la agrupación y clasificación con Weka
- Los tipos de delitos

2.4 Pruebas y verificación de resultados

En esta fase se realizó todas las actividades que plantearon; además se generaron los modelos con la ayuda de las plataformas de Weka y RStudio.

Lo primero que se hizo fue pasar la base de datos a Weka (ver figura 3), de ahí se pudo observar los campos ya seleccionados para realizar la investigación.

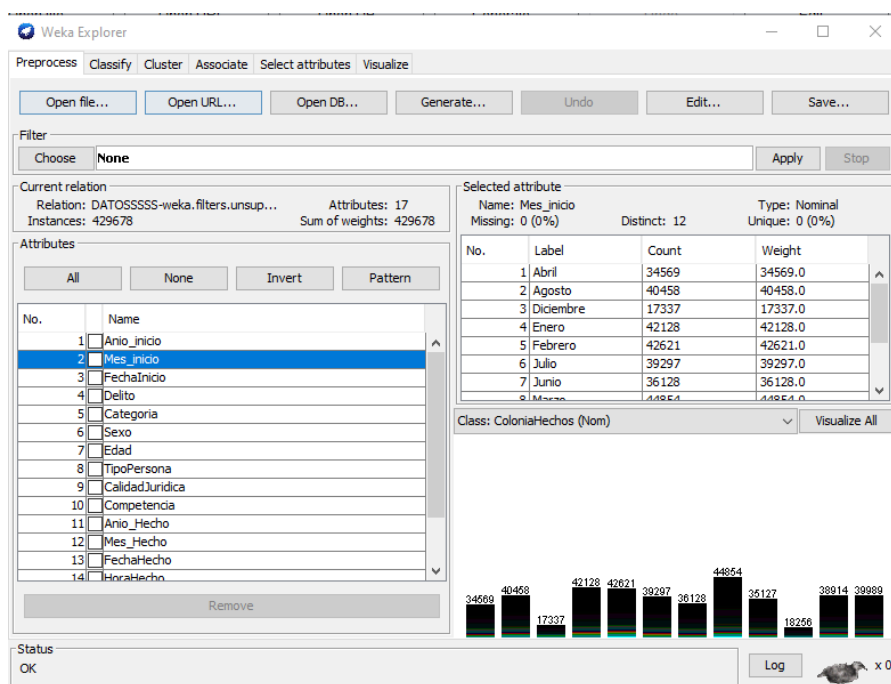


Ilustración 3 Campos seleccionados para trabajar en Weka

Se comenzó a realizar el preproceso para ello se va a utilizar un filtro a los campos, se seleccionó el filtro de los Supervisados, la opción de atributos y se seleccionó Discretizado la cual convierte los atributos numéricos en nominales, ya que es necesario para la clasificación y por último se le aplico a cada uno de los campos. Todo lo anterior nos ayudó para la construcción del árbol de decisión, los campos que se utilizaron fueron sexo, categoría, edad, Mes_Hecho, HoraHecho y AlcaldíaHechos, de las 429678 instancias (registros) que formaba nuestra base, Weka tomó 253445 instancias (58.9849%) para construir el modelo; además de que se utilizó la Cross-validation con un numero de particiones de 10 y la variable categórica que fue seleccionada para la predicción fue el sexo.

Se probó el segundo modelo los k vecinos más cercanos (SimpleKMeans), en la opción de Cluster Mode se utilizó Use training set para que tome todos los datos disponibles y aplicarlos otra vez sobre los mismos; se configuro para obtener 3 grupos(clústeres); se obtuvo 3 iteraciones; en el modelo y evaluación del conjunto de entrenamiento se obtuvo:

Instancias agrupadas:

0	204586 (48%)	1	149962 (35%)	2	75130 (17%)
---	---------------	---	---------------	---	--------------

El ultimo algoritmo utilizado en Weka fue las Reglas de Asociación, en la cual se utilizó el método A priori, el cual nos permitió buscar asociaciones entre los datos; en las configuraciones se colocó que se obtuvieran 10 reglas; el método mostro el total de instancias de 171871, con la confianza del 0.9.

En la parte de RStudio se trabajó la parte de la creación de tablas y las gráficas; además se utilizó para construir preguntas que se podían obtener de la base; preguntas como: ¿Cuál es el top 10 de delitos de la ciudad de México?, obtener las horas en las que más se presentaban los delitos, las colonias con más delitos registrados, etc.

3. RESULTADOS

Los resultados obtenidos a partir de la realización de cada algoritmo o técnica empleada:

3.1 Árbol de decisión

- Correspondiente a la creación del árbol de decisión se tomó el método tres.J48, en el cual se obtuvo el siguiente árbol:

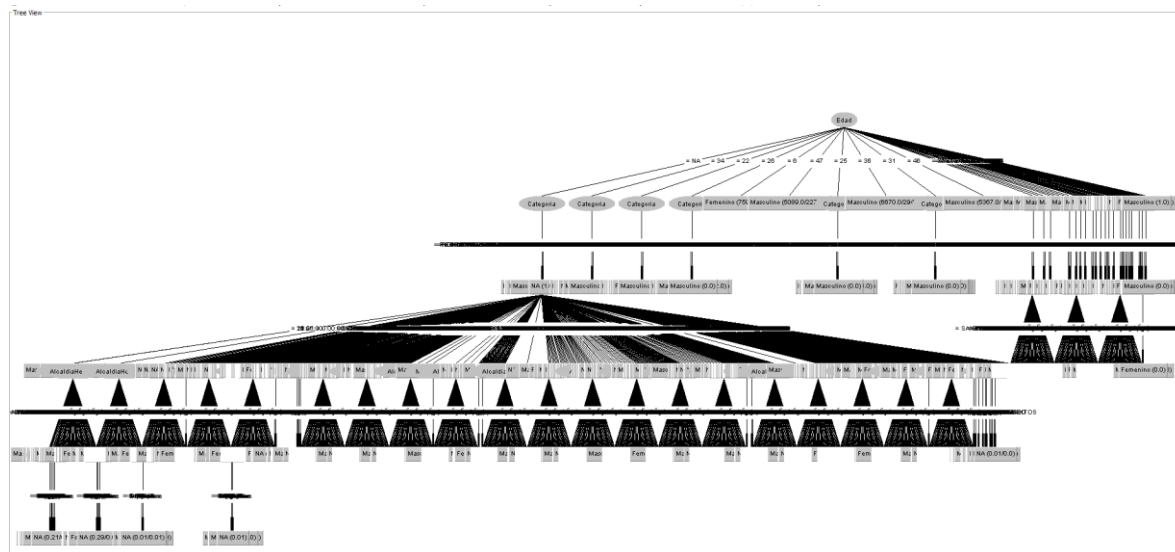


Ilustración 4 Arbol de decisión sobre la edad y su posible tipo de delito.

El método trabajo con un total de 429678 instancias, tomo todos los atributos; el árbol tomo un número de hojas de 11045 y el tamaño del árbol es de 11128 (véase ilustración 4); Instancias correctamente clasificadas 253445 (58.9849%) e instancias clasificadas incorrectamente 1762233 (41.0151%); cómo podemos notar el árbol está muy grande y debido a eso no se puede observar con claridad; el árbol nos explica:

- Si la categoría es Delito de bajo impacto, la hora esta entre las 00:00:00, la alcaldía es Iztapalapa, el sexo con mayor peligro es Masculino.
- Otro ejemplo es edad= 37 años, categoría= delito de bajo impacto, alcaldía de los hechos es igual a Gustavo a Madero el sexo con mayor peligro es Femenino.

Por último, se nos presenta la Matriz de confusión:

a	b	c	← classified as
78583	11525	390	a = NA
23101	122818	42326	b = Masculino
16117	82774	52044	c = Femenino

La cual nos dice que B se toma el sexo Masculino y C para femenino, tenemos que en B de los 122818 registros que se analizaron 82774 correspondían al sexo Femenino, y así sucesivamente podemos analizar. Se tiene que tener en cuenta que la diagonal de ser mayor para comprobar la confiabilidad y como observar es muy confiable, ya que tomando B y C nada más analizamos la diagonal y se comprueba la confiabilidad.

3.2 Clúster

En la parte del agrupamiento Weka analizo un total de 429678 instancias (registros), con el total de 11 atributos que son: Anio_inicio, Mes_inicio, Delito Categoría, Sexo, Edad, Anio_Hecho, Mes_Hecho, HoraHecho, AlcaldiaHechos y ColoniaHechos. El modo de cluster fue el que usa todos los registros (Use training set), el resultado fue:

Final cluster centroids:

Attribute	Full Data (429678.0)	Cluster# 0 (204586.0)	1 (149962.0)	2 (75130.0)
i¿Anio_inicio	'All'	'All'	'All'	'All'
Mes_inicio	Marzo	Marzo	Enero	Julio
Delito	VIOLENCIA FAMILIAR	VIOLENCIA FAMILIAR	VIOLENCIA FAMILIAR ROBO A NEGOCIO SIN VIOLENCIA	
Categoría	DELITO DE BAJO IMPACTO	DELITO DE BAJO IMPACTO	DELITO DE BAJO IMPACTO	DELITO DE BAJO IMPACTO
Sexo	Masculino	Femenino	Masculino	Masculino
Edad	NA	NA	NA	NA
Anio_Hecho	'All'	'All'	'All'	'All'
Mes_Hecho	Marzo	Marzo	Enero	Julio
Mes_Hecho	12:00:00	12:00:00	12:00:00	12:00:00
HoraHecho	IZTAPALAPA	CUAUHTEMOC	CUAUHTEMOC	IZTAPALAPA
AlcaldiaHechos	NA	NA	NA	NA
ColoniaHechos				

Ilustración 5 Resultado de la tecnica de Clustering en Weka

El cual nos dice que se realizaron 3 iteraciones, en el cluster 0 se encontro todos los años , el mes fue Marzo el cual registro y predominio la Violencia familiar en mujeres, no se registro la edad y los hechos se registraron mayormente en la alcaldia Cuauhtemoc; en el cluster 1 tenemos que el delito que mas domino fue violencia familiar pero en este caso fue al sexo Masculino, pero los hechos ocurrieron mayormente en el mes de enero, de igual forma en la alcaldia Cuauhtemoc; por ultimo tenemos que en el cluster 2 el delito que mas predominio fue robo a negocio sin violencia, las victimas fueron la mayoria Masculinos, en el mes de julio en la alcaldia Iztapalapa.

El total de instancias evaluadas por cada cluster fue de :

0 204586 (48%)

1 149962 (35%)

2 75130 (17%)

3.3 Reglas de Asociación

Por último se realizó la última técnica con la cual se encontraron 10 características que nosotros a simple vista no podríamos notar; las cuales a continuación se muestran 3 de ellas:

1. CalidadJuridica=VICTIMA Y DENUNCIANTE Competencia=FUERO COMUN 257697 ==> TipoPersona=FISICA 256210 <conf:(0.99)> lift:(1.26) lev:(0.12) [53580] conv:(37.01)

En la primera regla encontrada tenemos que si la calidad jurídica es de víctima y denunciante encontramos la competencia de fuero común es decir delitos que solo ocurren y se denuncian dentro de la Ciudad de México, con la confianza del 0.99.

2. CalidadJuridica=VICTIMA Y DENUNCIANTE 259911 ==> TipoPersona=FISICA 258407 <conf:(0.99)> lift:(1.26) lev:(0.13) [54037] conv:(36.9)

La siguiente regla dice que si la calidad es víctima y denunciante tenemos que las personas son de tipo física es decir personas individuales. La confianza es del 0.99 casi el 100%

3. Categoria=DELITO DE BAJO IMPACTO TipoPersona=FISICA 265104 ==> Competencia=FUERO COMUN 262709 <conf:(0.99)> lift:(1.01) lev:(0.01) [3589] conv:(2.5)

Nos dice que si la categoría es de bajo impacto y el tipo de persona es física tendremos delitos que solo ocurren y se denuncian dentro de la Ciudad de México.

3.4 Resultados con RStudio

Trabajando con R se realizaron algunas graficas como se muestran a continuación:



Ilustración 7 Gráfica sobre la comparación del número de carpetas abiertas por cada año (2019-2020)

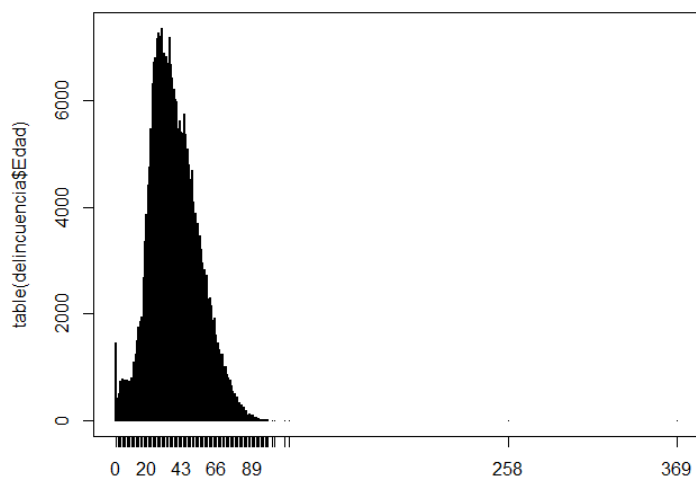


Ilustración 6 Grafica que muestra las edades entre las cuales se registró una mayor cantidad de delitos de los años 2019 y 2020

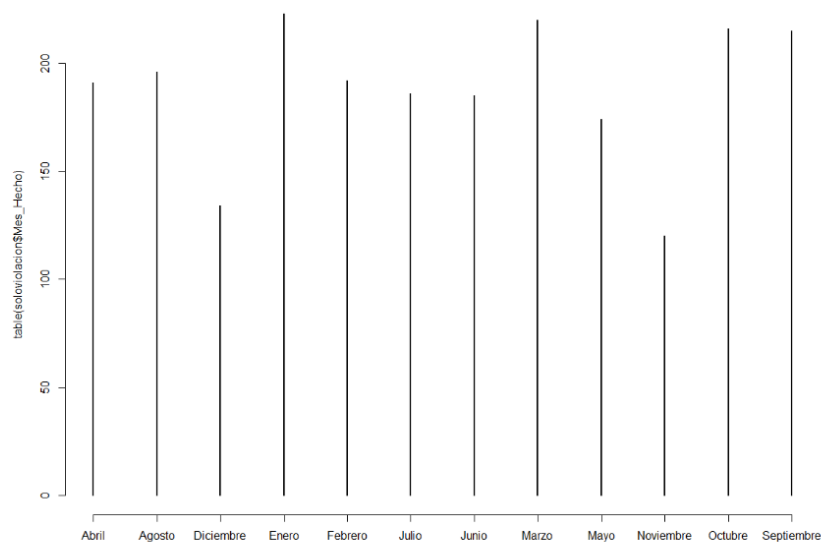


Ilustración 8 Gráfica de los meses y la cantidad de delitos registrados en cada uno.

4. CONCLUSIONES

Con todo el trabajo realizado de minería de datos en las dos herramientas que se mencionaron podemos concluir que:

- El año 2019 y el mes de marzo la FGJ registro el mayor número de denuncias para iniciar carpetas de investigación;
- Las horas en las que se registraron menos carpetas estuvieron entre: 0:00:05, 0:00:06 ,0:00:14, 0:00:15 ,0:00:16 y 0:00:23.
- El top 10 de los delitos más registrados son:

[1] "VIOLENCIA FAMILIAR"

[2] "FRAUDE"

[3] "AMENAZAS"

[4] "ROBO A NEGOCIO SIN VIOLENCIA"

[5] "ROBO A TRANSEUNTE EN VIA PUBLICA CON VIOLENCIA"

[6] "ROBO DE OBJETOS"

[7] "ROBO A NEGOCIO CON VIOLENCIA"

[8] "ROBO DE ACCESORIOS DE AUTO"

[9] "ROBO DE VEHICULO DE SERVICIO PARTICULAR SIN VIOLENCIA"

[10] "ROBO DE OBJETOS DEL INTERIOR DE UN VEHICULO"

- Las categorías más registradas son:

[1] "DELITO DE BAJO IMPACTO"

[2] "ROBO A TRANSEUNTE EN VIA PUBLICA CON Y SIN VIOLENCIA"


[3] "ROBO DE VEHICULO CON Y SIN VIOLENCIA"


[4] "ROBO A NEGOCIO CON VIOLENCIA"


[5] "ROBO A REPARTIDOR CON Y SIN VIOLENCIA"

- Las personas del sexo Femenino fueron las que presentaron un mayor número de delitos y apertura de carpetas de investigación, entre la edad de los 20 y 30 años.
- El tipo de persona que tuvo más registros fue de tipo Física, la calidad jurídica más encontrada fue de tipo víctima y denunciante, los delitos la mayoría fueron ocurridos y denunciados dentro de la ciudad.
- En la parte de los hechos se tiene que: marzo registro el mayor número de delitos, entre las horas 16:00:00, 11:00:00 ,14:00:00 ,15:00:00 ,10:00:00 y 12:00:00;
- El top 10 de las alcaldías con el mayor registro de delitos en los años 2019 y 2020 fueron: "IZTAPALAPA", "CUAUHTEMOC", "GUSTAVO A MADERO", "BENITO JUAREZ", "ALVARO OBREGON", "MIGUEL HIDALGO", "COYOACAN", "VENUSTIANO CARRANZA", "TLALPAN" y "AZCAPOTZALCO".
- Por último, se encontró el top 10 de las colonias con un alto número de delitos registrados en la ciudad: "CENTRO", "DOCTORES", "DEL VALLE CENTRO", "MORELOS", "ROMANORTE", "BUENAVISTA", "NARVARTE", "AGRICOLA ORIENTAL", "POLANCO" y "JUAREZ".



Hoy en la actualidad la ciudad de México cuenta con muchos trabajos de investigación de este tipo, donde se estudia lo antes hecho en este documento; pero ¿Qué pasa con estas investigaciones en las cuales describen esta problemática? ¿qué están haciendo las autoridades competentes para erradicar estos delitos?; como se puede ver en la ilustración 9 las personas que viven en la ciudad se encuentran molestos con este tipo de trabajos de investigación ya que no se hace nada para arreglar la situación, ya que solo se entregan resultados de investigaciones, pero no se entregan resultados sobre la erradicación de los delitos.


 Una vez mas la pregunta es: ¿y que va a hacer la "autoridad" con este "trabajo de inteligencia", ¿solo nos avisan para que no pasemos por ahí o harán su trabajo de erradicar la delincuencia? Habrían también de hacer un mapa de ministerios públicos y autoridades corruptas, por que de otro modo no se explica como es que la delincuencia aumenta si las autoridades trabajan tan arduamente.

Me gusta · Responder ·  9 · 45 sem

 Y todas las alcaldías

Me gusta · Responder · 45 sem

 Felicidades!!! descubrieron el hilo negro ..Y TALVEZ LES FALTO 2 o 3 DIGITOS ..Ahora  a trabajar q para eso les paga,, jijj

Me gusta · Responder ·  4 · 45 sem

 Y las autoridades de complices... tienen todos los datos y no hacen nada!!!

Me gusta · Responder ·  5 · 45 sem


 Y porqué no se ha doblado la vigilancia policiaca en las zonas ya ubicadas?

Ilustración 9 Comentarios de personas molestas por investigaciones que se realizan en la ciudad de México.

(UNIVERSAL, 2020)

A las conclusiones a las que llegue yo personalmente, es que la minería de datos es de gran ayuda ya que nos proporciona información de valor para determinar algunos tipos de situaciones o determinar factores de riesgo de manera oportuna; sin embargo, es necesario no solo quedarse con la información guardada es necesario su difusión para erradicar o decrementar factores de riesgo.

REFERENCIAS

- Corso, C. L. (2009). Aplicación de algoritmos de clasificación supervisada usando Weka . *Universidad Tecnológica Nacional, Facultad Regional Córdoba*, 1-11.
- INEGI. (2018). *Instituto Nacional de Estadística y Geografía*. Obtenido de <https://www.inegi.org.mx/temas/incidencia/>
- México, G. d. (s.f.). *Portal de datos de la ciudad de México*.
- México, G. d. (s.f.). *Portal de datos de la ciudad de México*. Obtenido de https://datos.cdmx.gob.mx/api/datasets/1.0/victimas-en-carpetas-de-investigacion-pgj/attachments/diccionario_de_victimas_actualizado_xlsx/
- México, G. d. (s.f.). *Portal de datos de la Ciudad de México*. Obtenido de <https://datos.cdmx.gob.mx/explore/dataset/victimas-en-carpetas-de-investigacion-pgj/information/>
- UNIVERSAL, E. (03 de marzo de 2020). *EL UNIVERSAL* . Obtenido de <https://www.eluniversal.com.mx/metropoli/estas-son-las-8-zonas-mas-peligrosas-de-la-cdmx>
- Valero Orea, S., Salvador Vargas, A., & Garcia Alonso, M. (2005). Minería de datos: predicción de la deserción escolar mediante el algoritmo de árboles de decisión y el algoritmo de los k vecinos más cercanos. 1-8.