



MINADO DE DATOS RESPECTO A LOS FACTORES IMPLICADOS EN LA INTERRUPCIÓN VOLUNTARIA DEL EMBARAZO DATA MINING REGARDING THE FACTORS INVOLVED IN THE VOLUNTARY INTERRUPTION OF PREGNANCY

Oswaldo Guerrero Navarrete

Estudiante de la Universidad Autónoma del Estado de México (UAEM), Centro Universitario
Nezahualcóyotl, Estado de México, México oguerreon001@alumno.uaemex.mx

Litzy Sarai López Pineda

Estudiante de la Universidad Autónoma del Estado de México (UAEM), Centro Universitario
Nezahualcóyotl, Estado de México, México

llopezp007@alumno.uaemex.mx

Nayeli Reyes Rosales

Estudiante de la Universidad Autónoma del Estado de México (UAEM)
Centro Universitario Nezahualcóyotl, Estado de México, México,

nreyesr001@alumno.uaemex.mx

Sergio David Ortiz García

Estudiante de la Universidad Autónoma del Estado de México (UAEM)
Centro Universitario Nezahualcóyotl, Estado de México, México,

sortizg001@alumno.uaemex.mx

Juan Carlos Santiago Araujo Estudiante de la
Centro Universitario Nezahualcóyotl, Estado de
México, México

jsantiagoa001@alumno.uaemex.mx

RESUMEN

La minería de datos tiene como propósito el procesar datos de un dominio para encontrar patrones de comportamiento útiles para la toma de decisiones. En este trabajo se presenta el proceso de minería de datos que se lleva a cabo mediante una investigación acerca del aborto. Cuyo propósito fue analizar los datos obtenidos en diferentes bases o repositorios de datos, para identificar los motivos o factores que presentan las mujeres para interrumpir el embarazo. Tomando las variables como año, edad, estado civil, ocupación, número de hijos, motivo de aborto y entidad. Después de todo este proceso se pudo determinar el principal factor para tomar la decisión del aborto, el cual es la interrupción voluntaria, es decir, que en ese

momento no desean ser madres, además de que el estado civil también es un factor característico.

Palabras Claves: Minería De Datos, Weka, Aborto, CRISP-DM, México.

ABSTRACT

The purpose of data mining is to process data from a domain to find patterns of behaviour useful for decision making. This paper presents the data mining process carried out through research on abortion. Its purpose was to analyse data obtained from different databases or repositories in order to identify the reasons or factors that women have for terminating a pregnancy. Variables such as year, age, marital status, occupation,



number of children, reason for abortion and entity were taken into account. After this process, it was possible to determine the main factor in the decision to have an abortion, which is voluntary termination, that is, that at that moment they do not wish to become mothers, in addition to the fact that marital status is also a characteristic factor.

Keywords: Data Mining, Weka, Abortion, CRISP-DM, Mexico.

Fecha de envío: 10/06/2022

Fecha de aprobación: 11/07/2022

Fecha de publicación: 01/01/2023

INTRODUCCIÓN

En los últimos años el tema del aborto en México ha incrementado su popularidad entre la población. ¿Pero por qué el aborto es un problema en nuestro país? Como es conocido, México ocupa uno de los primeros lugares en cuanto a población se refiere, concentrándose mayormente en la Ciudad de México con 9,209,944 personas tan solo en 2020 según el INEGI (Instituto Nacional de Geografía y Estadística), esto debido a múltiples factores como son: La falta de educación sexual, escasos de empleos en el lugar de origen, centralización en la CDMX y penalización del aborto.

Según la nueva sociedad en la despenalización del aborto en México fue hasta abril de 2007 que la Ciudad de México aprobó la despenalización y comenzó a ofrecer el servicio en hospitales públicos e instituciones de salud. El artículo argumenta que esto fue posible gracias al impulso de los grupos feministas, la polarización electoral y el triunfo del Partido de la

Revolución Democrática (PRD) en el entonces Distrito Federal.

Cabe mencionar que de acuerdo con datos proporcionados por el Sistema de Información de Interrupción Legal del Embarazo, tan solo durante el año 2021 se llevaron a cabo un total de 9,015 intervenciones registradas apenas en la Ciudad de México, que si bien es menor al número registrado en años anteriores, aun así, sigue siendo una cantidad alarmante, lo que resalta la importancia de saber cuáles exactamente son los factores que influyen en este proceso, de tal modo que quizá se pueda llegar a la raíz de la situación. Otro factor importante que gira alrededor del aborto es la cantidad de fallecimientos debido a los abortos clandestinos que se llevaron a cabo antes de la legalización y regularización de este. De igual manera, de acuerdo con el artículo “La mortalidad materna y el aborto en México” escrito por la Dra. Sonia B. Fernández y sus compañeros Gonzalo Gutiérrez y Ricardo Viguri en el año 2012, “... se enfatiza el aborto como causa desencadenante. No debe dejarse de lado que la muerte materna es uno de los factores que mayormente afecta la salud biopsicosocial de los hijos huérfanos sobrevivientes.” Considerando que, “La mortalidad materna se define como la muerte de una mujer durante el embarazo, el parto o dentro de los 42 días después de su terminación, por cualquier causa relacionada o agravada por el embarazo, el parto, el puerperio o su manejo, pero no por causas accidentales.” Pues curiosamente, si bien el número de muertes maternas ha descendido de 1477 a 992 defunciones durante el periodo 1990-2010, la razón de mortalidad materna ha permanecido



prácticamente constante, con una ligera tendencia a la baja durante los años más recientes.

El motivo del trabajo es la recolección de datos para un análisis más concreto de las condiciones o motivos que tiene una mujer embarazada quiera realizar un aborto además de que las expectativas en torno a este trabajo son demostrar el por qué una mujer embarazada aborta, ya que hay muchos aspectos a considerar tales como los de no tener los recursos necesarios para ofrecerle a un recién nacido, ya sea porque son menores de edad, mujeres solteras o en algunos casos sea por una situación de abuso sexual. Muchas mujeres también suelen tomar la decisión de realizarse un aborto porque lo ven como una salida más cómoda para poder seguir su vida normal.

Tal y como se puede apreciar, el camino es largo de recorrer y hay mucho que investigar. Con este trabajo se podrá dar a conocer un mejor análisis de como los motivos de una mujer embarazada practican el aborto sin tener en cuenta algún aspecto moral o político, ya que este último no afecta a un grupo político. Por ello se tratará de abordar la situación de la manera más neutra posible, inhibiendo cualquier opinión que cualquiera de los involucrados pueda llegar a tener, de modo que se pueda ofrecer una investigación objetiva e imparcial. Es importante señalar dos aspectos, el primero es que tal y como se mencionó no se verá reflejada ninguna opinión, ni ninguna influencia por parte de terceros, por lo que en todo momento deberá tenerse presente el objetivo informativo y estadístico que se presenta. Lo segundo, es que, parte del análisis que se presentará, puntuando

en aquel donde los números y estadísticas no están tan presentes sino hasta el momento de evaluar resultados, serán meras conjeturas, ya que ninguno de los involucrados está especializado en el tema, por lo que ninguno tiene la autoridad de afirmar o negar un beneficio o un prejuicio más allá del interpretado matemáticamente. Una vez aclarado el punto, se menciona que, el análisis realizado contara con el uso de las herramientas tecnológicas con las que se cuenta al momento en que se redacta este documento.

MÉTODOS Y MATERIALES

Se realizó un estudio en las instalaciones del Centro Universitario UAEM Nezahualcóyotl, específicamente en el laboratorio de minería de datos del año en curso, acerca de las causas que propician la decisión del aborto en las entidades de la República Mexicana, mediante la búsqueda en diferentes bases de datos, entre las cuales destacan: INEGI, Datos Abiertos De México, Explore-Datos CDMX y Catálogo Nacional De Indicadores, basado en un período de 2016 a 2019. Cabe mencionar que los términos y criterios de búsqueda durante la recopilación de datos fueron aquellos datos que nos indiquen que fue una interrupción legal del embarazo y de estos mismos solo se fueron recopilando los que muestren los posibles factores que determinen el por qué se lleva a cabo este proceso.

Para la clasificación de los datos se utilizó el software Weka como herramienta para el proceso de minería de datos. Weka se caracteriza por tener gran variedad de opciones (Kuri,2020),



pero para este estudio se utilizaron los pertenecientes a clasificadores, específicamente árboles de decisión, donde se seleccionaron tres algoritmos que son J48, Random Forest y RepTree, los cuales presentan mayor precisión en los resultados.

También existen diversas técnicas y metodologías de minería de datos que se pueden utilizar y que pueden ser adecuadas para cada caso en concreto, pero en el presente proyecto se desarrolla la metodología CRISP-DM (Cross Industry Standard Process for Data Mining) conocida como el estándar de factor para el desarrollo de la minería de datos y los proyectos de descubrimiento de conocimiento (Marbán,2018). La metodología CRISP-DM es capaz de estructurar el ciclo de vida de un proyecto en seis fases, que interactúan entre ellas de forma iterativa durante el desarrollo.

La metodología CRISP-DM establece un proyecto de minería de datos como una secuencia de fases:

1. Comprensión del negocio.
2. Compresión de los datos.
3. Preparación de los datos.
4. Modelado.
5. Evaluación.
6. Despliegue.

DESARROLLO

Para realizar el estudio, se utilizaron modelos de Minería De Datos capaces de realizar predicciones en un entorno real utilizando datos reales. Se aplicó la metodología CRIPS-DM ya que es una metodología neutral, mejora la calidad de resultados en los proyectos, debido a que está estructurada y ha sido ampliamente adoptada en análisis

de minería de datos (Betancor,2019; Harden y Ogden,1999).

El universo de estudio está compuesto por 19,870 registros de abortos. Debido a ello, fue necesario usar el software WEKA para llevar a cabo el análisis y construcción de cada uno de los modelos de minería de datos.

SELECCIÓN DE DATOS

Con el propósito de llevar a cabo este estudio, los datos seleccionados provienen de las bases descritas anteriormente en la sección de Materiales y Métodos. El criterio de selección aplicado fue lo suficientemente amplio para permitir incluir datos de relevancia en función de los objetivos del proyecto. Debido a esto a continuación se explican los criterios que se llevaron a cabo para la selección de datos.

VARIABLES DE ESTUDIO

Finalmente las variables más representativas para abordar la problemática fueron: *Año, Estado Civil, edad, nivel de educación, ocupación, entidad, número de hijos y motivo del aborto*

PRE-PROCESAMIENTO DE DATOS

Durante el desarrollo de esta etapa se realizó una limpieza de datos, lo cual mejorará el estado de la base, permitiendo obtener patrones de mejor calidad y mayor precisión en los resultados. De acuerdo con esto, es importante mencionar que por motivos de datos incompletos y datos anómalos (aquellos que contenían "N/A") se



tuvieron que eliminar algunos registros. A continuación, se describen cada una de los Cuadros agregadas a este proyecto. En el Cuadro 1, podemos observar los atributos seleccionados, los cuales son considerados los más importantes para desarrollar este estudio de minería de datos.

Cuadro 1. Explicación de los atributos seleccionados

Atributo	Definición	Descripción
AÑO	Año	Período en el que se registró el aborto.
EC	Estado Civil	Presenta información del estado civil de la mujer
ED	Edad	Indica el número de años cumplidos
NED	Nivel de educación	Grado de estudios más alto que tiene la persona.
OC P	Ocupación	Indica el empleo o trabajo que tiene la persona.
ENT	Entidad	Residencia de la mujer que decide abortar
NH	Número De Hijos	Información de número de hijos.
MT A	Motivo Del Aborto	Motivo para la interrupción del embarazo.

Cuadro 2. Descripción de los atributos asociados al estado civil

Atributo	Definición
CAS	Casada
SOL	Soltera
DIV	Divorciada
ULI	Unión L

Cuadro 3. Descripción de los atributos asociados del nivel de educación.

Atributo	Definición
PRIM	Primaria
SEC	Secundaria
PREP	Preparatoria
LIC	Licenciatura
MTRI	Maestría

Cuadro 4. Descripción de los atributos asociados a ocupación

Atributo	Definición
EMP	Empleada
DES	Desempleada
EST	Estudiante

Cuadro 5. Descripción de los atributos asociados a motivo de aborto



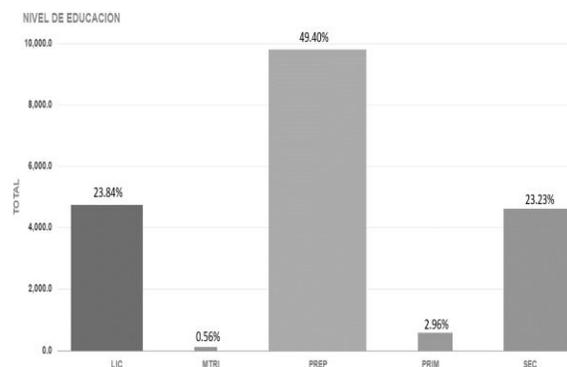
Atributo	Definición
INTV	Interrupción Voluntaria
FDM	Falla Del Método
PRDS	Problemas De Salud
PRV	Proyecto De Vida
SECO	Situación Económica
VIO	Violación

ANÁLISIS DE DATOS

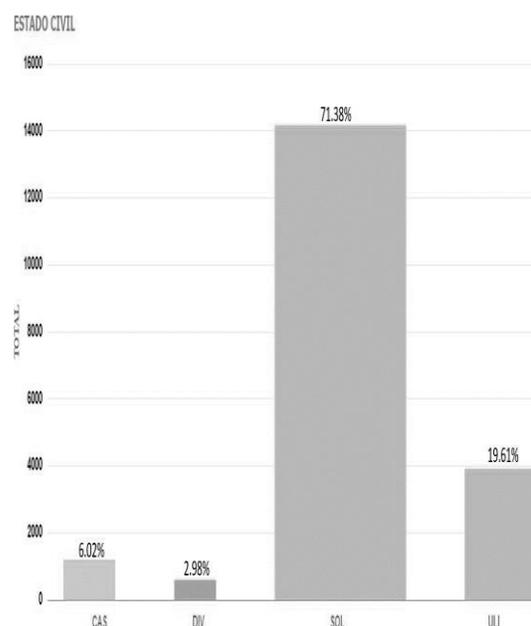
Una vez que se han descrito los datos, se procede a analizarlos, esto implica aplicar diferentes técnicas que revelarán las propiedades de los datos y así determinar su consistencia (De la Fuente, 2015; Fernández et.al, 2014; ACAI, 2014). Para entender los patrones obtenidos gracias a la minería de datos se requieren métodos como la agrupación (clustering) o reglas de asociación, entre otras (Figuroa et.al, 2012; González et.al, 2002). La finalidad de esta etapa es descubrir una estructura general para los datos, mediante la creación de Cuadros de frecuencia y gráficos de distribución de los datos. Para el desarrollo de esta etapa se utilizó el software Oracle Data Visualization [Forbes, 2020], el cual permitió observar el comportamiento de los datos de forma gráfica para identificar patrones, tendencias y correlaciones.

La Gráfica 1, muestra el análisis de frecuencia de la variable Estado Civil. Podemos observar que el estado civil es un factor característico en el tema del aborto, ya que las mujeres que deciden interrumpir el embarazo generalmente son solteras (Fernández et.al, 2012,

García y Seuc,1997; González et.al 2008; González et.al, 2013), con un alto porcentaje del 71.38%.



Gráfica 1. Análisis de Frecuencia de los Registros de Aborto mediante el Estado Civil

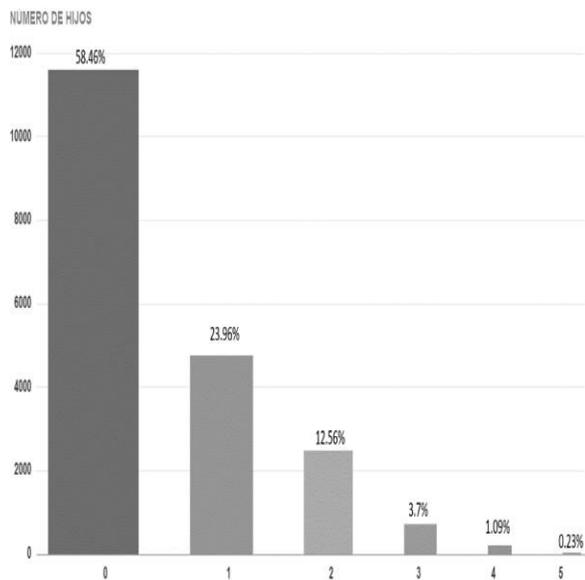


Fuente: Elaboración Propia (2022)

Posteriormente se muestra la Gráfica 2, donde se encuentra el análisis de la variable edad, presentando mayor número de registros sobre aborto las mujeres que tienen 20 años (Castañeda et.al, 1999; González et.al, 2008; González et.al, 2013).



Gráfica 2. Análisis de Frecuencia de

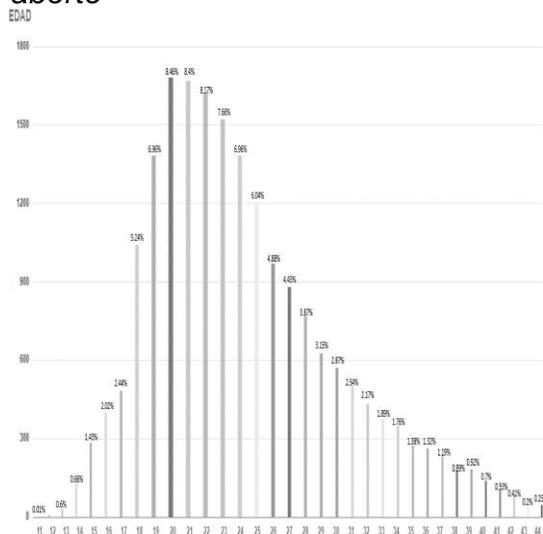


Registros de Aborto por Edad

Fuente: Elaboración Propia (2022)

La Gráfica 3, presenta el registro de la variable sobre el nivel de educación, donde observamos que las mujeres que deciden abortar tienen una formación suficiente, es decir, 49.40% tiene preparatoria [Islas, 2008].

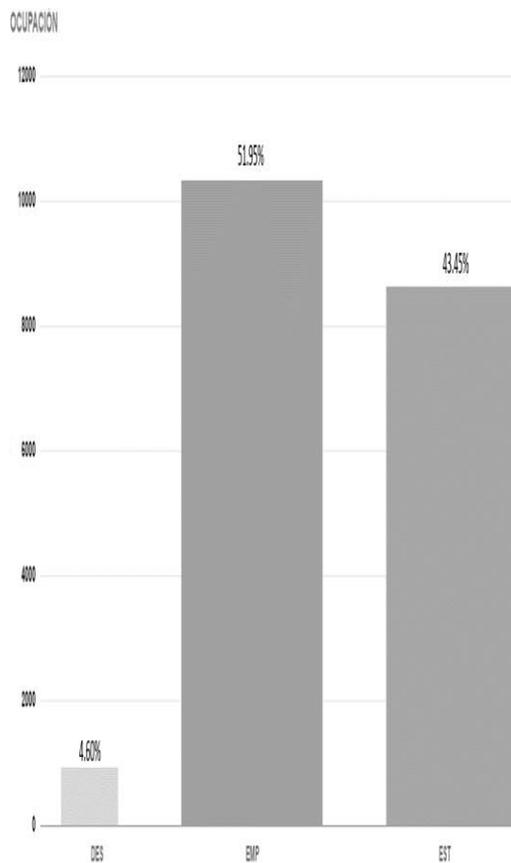
Gráfica 3. Análisis de Frecuencia del nivel de educación en registros sobre el aborto



Fuente: Elaboración Propia (2022)

La Gráfica 4 muestra que generalmente las mujeres con un empleo [Cabezas et.al, 1998] tienen mayor posibilidad de decidir interrumpir su embarazo, pues el porcentaje es 51.95%.

Gráfica 4. Análisis de Frecuencia de Registros de Aborto por Ocupación

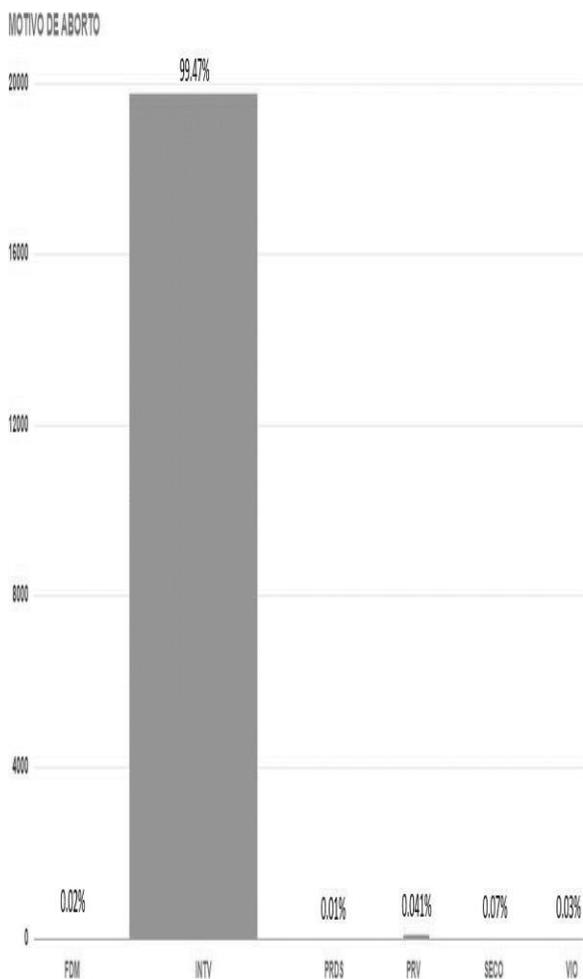


Fuente: Elaboración Propia (2022)

La Gráfica 5, nos muestra un análisis del número de hijos que posiblemente podría tener una mujer que decide abortar, pero podemos darnos cuenta que generalmente las mujeres sin hijos son las que presentan mayor registro de aborto con un 58.46%.



Gráfica 5. Análisis de Frecuencia de Registros de Aborto por Número de Hijos



Fuente: Elaboración Propia (2022)

Finalmente, la Gráfica 6, observamos que la interrupción voluntaria del embarazo tuvo mayor incidencia [Data Ming, 2018], registrando un porcentaje del 99.47% lo cual es un registro demasiado alto.

Gráfica 6. Análisis de Frecuencia de los Principales motivos del Aborto

Correctly Classified Instances	19753	99.4112 %							
Incorrectly Classified Instances	117	0.5888 %							
Mappa statistic	0.0076								
Mean absolute error	0.0035								
Root mean squared error	0.0445								
Relative absolute error	96.4155 %								
Root relative squared error	106.3395 %								
Total Number of Instances	19870								
=== Detailed Accuracy By Class ===									
	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.999	0.990	0.995	0.999	0.997	0.025	0.639	0.996	INTV
	0.000	0.000	0.000	0.000	0.000	-0.000	0.601	0.002	SECO
	0.000	0.000	0.000	0.000	0.000	-0.001	0.593	0.012	PRV
	0.000	0.000	0.000	0.000	0.000	-0.000	0.494	0.000	FEM
	0.000	0.000	?	0.000	?	?	0.573	0.001	VIO
	0.000	0.000	?	0.000	?	?	0.496	0.000	PRDS
Weighted Avg.	0.994	0.985	?	0.994	?	?	0.638	0.991	
=== Confusion Matrix ===									
	a	b	c	d	e	f	<-- classified as		
19753	3	8	1	0	0	0	a = INTV		
12	0	0	1	0	0	0	b = SECO		
81	0	0	0	0	0	0	c = PRV		
3	0	0	0	0	0	0	d = FEM		
6	0	0	0	0	0	0	e = VIO		
2	0	0	0	0	0	0	f = PRDS		

Fuente: Elaboración Propia (2022)

CLASIFICACIÓN

El modelo para un árbol de clasificación se presenta como una estructura jerárquica para mostrar y establecer las relaciones entre la variable de dependiente y el conjunto de variables predictoras (Langer et.al, 1998).

Como podemos observar en la Figura 1, 2 y 3, correspondientes a cada uno de los algoritmos mencionados anteriormente podemos visualizar el análisis por Motivo De Aborto, donde se representan que en los registros para el período 2016- 2019, la principal causa para abortar fue la interrupción voluntaria mostrando el 99%. Cabe mencionar que en los algoritmos para el análisis se obtuvo el mismo resultado.



Figura 1. Clasificación de Registros de Aborto por MTA usando J48

Fuente: Elaboración Propia (2022)

```

Correctly Classified Instances 19765 99.4716 %
Incorrectly Classified Instances 105 0.5284 %
Kappa statistic 0
Mean absolute error 0.0035
Root mean squared error 0.0419
Relative absolute error 97.4446 %
Root relative squared error 99.9996 %
Total Number of Instances 19870

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.995 1.000 0.997 ? 0.488 0.995 INTV
0.000 0.000 ? 0.000 ? ? 0.419 0.001 SECO
0.000 0.000 ? 0.000 ? ? 0.494 0.004 PRV
0.000 0.000 ? 0.000 ? ? 0.150 0.000 FFM
0.000 0.000 ? 0.000 ? ? 0.300 0.000 VIO
0.000 0.000 ? 0.000 ? ? 0.100 0.000 PRDS
Weighted Avg. 0.995 0.995 ? 0.995 ? ? 0.488 0.989

=== Confusion Matrix ===

a b c d e f <-- classified as
19765 0 0 0 0 0 0 | a = INTV
13 0 0 0 0 0 0 | b = SECO
81 0 0 0 0 0 0 | c = PRV
3 0 0 0 0 0 0 | d = FFM
6 0 0 0 0 0 0 | e = VIO
2 0 0 0 0 0 0 | f = PRDS
  
```

Figura 2. Clasificación de Registros de Aborto por MTA usando RandomForest

Fuente: Elaboración Propia (2022)

```

Correctly Classified Instances 19765 99.4716 %
Incorrectly Classified Instances 105 0.5284 %
Kappa statistic 0
Mean absolute error 0.0035
Root mean squared error 0.0419
Relative absolute error 97.4446 %
Root relative squared error 99.9996 %
Total Number of Instances 19870

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.995 1.000 0.997 ? 0.488 0.995 INTV
0.000 0.000 ? 0.000 ? ? 0.419 0.001 SECO
0.000 0.000 ? 0.000 ? ? 0.494 0.004 PRV
0.000 0.000 ? 0.000 ? ? 0.150 0.000 FFM
0.000 0.000 ? 0.000 ? ? 0.300 0.000 VIO
0.000 0.000 ? 0.000 ? ? 0.100 0.000 PRDS
Weighted Avg. 0.995 0.995 ? 0.995 ? ? 0.488 0.989

=== Confusion Matrix ===

a b c d e f <-- classified as
19765 0 0 0 0 0 0 | a = INTV
13 0 0 0 0 0 0 | b = SECO
81 0 0 0 0 0 0 | c = PRV
3 0 0 0 0 0 0 | d = FFM
6 0 0 0 0 0 0 | e = VIO
2 0 0 0 0 0 0 | f = PRDS
  
```

Figura 3. Clasificación de Registros de Aborto por MTA usando RepTree

Fuente: Elaboración Propia (2022)

La segunda variable de mayor peso fue el Estado Civil, para determinar esto se usaron dos algoritmos de árboles de decisión, entre los cuales están J48, y Hoeffding tree; y uno de reglas de decisión llamado JRip. Cada uno de estos algoritmos se seleccionó debido a que alcanzaron mejores resultados y cabe resaltar que para las opciones de prueba se utilizó el método de validación cruzada.

En la ejecución de cada algoritmo se dio como resultado que el atributo de mayor incidencia fue “solteras”, cada uno con su respectivo porcentaje. Seguido en la Figura 4, muestra el algoritmo Hoeffding tree con un porcentaje de 71%. Finalmente, en la Figura 5, se encuentra el análisis de esta variable, usando el algoritmo JRip, donde se presenta un porcentaje de 99%.

```

Correctly Classified Instances 14132 71.1223 %
Incorrectly Classified Instances 5738 28.8777 %
Kappa statistic 0.093
Mean absolute error 0.2014
Root mean squared error 0.323
Relative absolute error 89.9891 %
Root relative squared error 96.5669 %
Total Number of Instances 19870

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.081 0.013 0.293 0.081 0.127 0.127 0.705 0.150 CAS
0.970 0.876 0.734 0.970 0.836 0.184 0.708 0.823 SOL
0.069 0.029 0.367 0.069 0.116 0.084 0.622 0.282 ULI
0.010 0.003 0.086 0.010 0.018 0.020 0.732 0.082 DIV
Weighted Avg. 0.711 0.632 0.616 0.711 0.628 0.156 0.691 0.654

=== Confusion Matrix ===

a b c d <-- classified as
97 975 107 18 | a = CAS
104 13762 293 25 | b = SOL
76 3532 267 21 | c = ULI
54 473 60 6 | d = DIV
  
```

Figura 4. Clasificación de Registros de Aborto por Estado Civil usando Hoeffding tree

Fuente: Elaboración Propia (2022)



Figura 5. Regla JRip para el atributo

Estado Civil

Correctly Classified Instances	14136	71.041 %
Incorrectly Classified Instances	5694	28.6059 %
kappa statistic	0.0139	
Mean absolute error	0.2229	
Roc mean squared error	0.3341	
Relative absolute error	99.6015 %	
Roc relative squared error	99.9526 %	
Total Number of Instances	19870	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.011	0.001	0.406	0.011	0.021	0.059	0.518	0.069	CMS
	0.996	0.995	0.716	0.996	0.933	0.061	0.507	0.717	SOL
	0.011	0.004	0.400	0.011	0.022	0.038	0.505	0.200	UOL
	0.000	0.000	?	0.000	?	?	0.511	0.035	OTV
Weighted Avg.	0.714	0.704	?	0.714	?	?	0.516	0.536	

=== Confusion Matrix ===

a	b	c	d	←-- classified as
13	1170	14	0	a = CMS
9	14129	46	0	b = SOL
6	3846	44	0	c = UOL
4	593	6	0	d = OTV

Fuente: Elaboración Propia (2022)

RESULTADOS

Para demostrar que la hipótesis que planteo al inicio del trabajo se presentaron los resultados que se obtuvieron de los 19870 registros de abortos que van del período del 2016 al 2019 en México. En el análisis se determinó que la variable con mayor incidencia es “MTA” (Motivo del aborto), teniendo en cuenta que el factor más representante es la interrupción voluntaria (INTV), (véase en la Figura 1, 2 y 3), lo cual demuestra que del total de registros, el 99% de los casos se presenta esta causa, por este sentido, deducimos que se llevan a cabo este procedimiento porque no quieren ser madres en ese momento de su vida, de manera que esta decisión la llegan a tomar de manera libre y sobre todo voluntaria a pesar de algunos riesgos que puedan tener.

Posteriormente, la siguiente variable que es considerada como un factor

determinante, es la de Estado Civil (véase en la Figura 4 y 5). Esta variable es tomada como importante para este análisis, pues por lo que se demuestra en los resultados las mujeres que no tienen pareja tienen más probabilidades de acudir a realizar el aborto, ya que en los 19870 registros que se obtuvieron, el 99% fueron mujeres solteras. Por lo tanto, en los resultados del análisis se observa que el mayor porcentaje de abortos apareció en las mujeres con este estado civil (solteras).

Mencionado lo anterior el no contar con una pareja o el ser soltera incrementa la posibilidad de una interrupción voluntaria del embarazo con respecto a mujeres que son casadas o están en unión libre. Las variables como edad y número de hijos, no fueron estadísticamente significativas ya que el número de hijos no se le considera un riesgo o factor para la mujer que quiere realizar el aborto, sin embargo la edad se puede tomar como una variable que pueda ayudar en un análisis más extenso, sin embargo en este estudio se determinó que las mujeres que abortan son de cualquier edad demostrando que esta variable no es de gran relevancia en este estudio y no tiene gran impacto para el problema que se planteó, pues los factores que son importantes siguen siendo el motivo del porque se llegan a realizar un aborto por lo que los resultados obtenidos con el proceso de minería de datos, permitieron alcanzar el objetivo propuesto de buscar factores para la interrupción del embarazo.

DISCUSIÓN

Tomando en cuenta el trabajo de (Gallego et.al, 2015) en *Aplicación de técnicas de minería de datos en atención primaria en salud (aps) para el*



análisis de riesgos en mujeres gestantes de la población manizaleña atendida por asbsbasalud, donde la razón de ser de las variables consideradas y el conjunto de datos que se trató, fue para identificar los factores de riesgo a los que se somete una mujer en etapa gestante que podrían desembocar en el aborto **involuntario** del producto del embarazo. Mientras que nuestros resultados e hipótesis se enfocaron principalmente en conocer las razones del porque una mujer decide abortar **voluntariamente**, y donde nuestros resultados contribuyen para poder enfocarse en un sector social para no implementar estrategias contra el aborto sino entender las necesidades de las mujeres en esta situación y facilitar el acceso al servicio contemplando todas las variables sociales existentes en nuestro entorno, dejando a consideración de cada uno, las aplicaciones de los resultados que desemboquen en una mayor aportación en este tema del aborto.

Aún con esta disparidad, es conveniente mencionar que los datos utilizados para ambos estudios fueron recolectados de diferentes fuentes a pesar de utilizar la misma metodología, el enfoque principal de los estudios cambia los resultados como si fueran temas de índole opuestas pese a que las bases de los dos sean la minería de datos junto con el aborto.

El aborto es el método más practicado para resolver embarazos indeseados, y su elevada incidencia constituye un problema de salud (Cabezas et al,1998), detrás de todo esto existen factores que propician esta decisión, durante el desarrollo de este proyecto se obtienen como resultados que la búsqueda de la

interrupción del embarazo está asociada con una decisión voluntaria y el estado civil.

Existen más artículos en donde además del estado civil se obtiene como factor predominante la edad, como el caso del trabajo de Pérez Santos donde analizó 75 expedientes clínicos de mujeres que solicitaron aborto inducido tardío de su embarazo en un hospital de la ciudad de La Habana y encontró que más de 60% eran menores de 24 años y solteras (1988). En otro estudio se evaluó la eficacia del misoprostol para interrumpir embarazos en el primer trimestre entre 141 mujeres notificando que el promedio de edad era de 23 años y que 52% eran solteras (Velasco et. al, 2008).

CONCLUSIONES

Controversial o no, la interrupción voluntaria del embarazo es un tema que sobre el que nuestra sociedad ha posado su mirada y lo ha convertido en un asunto interés público. Se puede decir en entonces que el sector de la población que más suele participar en esta práctica son mujeres solteras en sus 20's incursionadas en el campo laboral, pero con tentativa de estudios inconclusos, lo cual apunta a que estas situaciones sean un derivado de la deficiente calidad educativa, así como la situación económica tan precaria no solo presente de manera individual sino a un nivel más general.

Ahora bien, es importante entender que la relación que guarda con otros elementos más allá de los aquí presentados, si bien es cierto, se ha dicho no se tiene la intención de indagar en tales áreas por el momento, aun así, eso no significa que no estén presentes,



por lo que lo aquí presentado es apenas la punta del iceberg.

Como dato adicional, ¿Es bueno o malo?, ¿Es correcto o incorrecto?, siendo realistas, ¿quién realmente dictamina esos parámetros?, la moral es relativa y lo que hoy es prohibido mañana será permitido y después obligatorio y viceversa. El debate esta sobre la mesa y aún queda mucho por ver, pero por lo pronto lo único seguro de afirmar es que independientemente de si la postura al respecto es a favor o en contra, lo ideal sería llegar a la raíz del asunto, averiguar los motivos auténticos detrás de y solo entonces adoptar una postura.

REFERENCIAS

- ACAI (2014). Condiciones que motivan el aborto. Asociación De Clínicas Acreditadas Para La Interrupción Del Embarazo. Recuperado el 27 de marzo de 2022, de <https://www.acaive.com/pdf/Condiciones-quemotivan-el-aborto-provocado-Estudio-ACAI-12-06-12.pdf>
- Belinchón Monjas, Y. (s. f.). Minería De Datos. Recuperado 26 de marzo de 2022, de <http://www.it.uc3m.es/jvillena/irc/practicas/10-11/15mem.pdf>
- Bernal Becerril M.L. (2010). Caracterización de la población usuaria que ingresa al programa interrupción legal del embarazo en un hospital del sector salud del D.F. *Enferm. univ.* vol.7, n.1, pp.8-15. ISSN 2395-8421. Disponible en: <http://www.scielo.org.mx/scielo.php?pid=S1665-70632010000100002&script=sci>
- Betancor, P. (2019, 16 enero). Introducción al machine learning con Weka. *Adictos al trabajo.* Disponible en: <https://www.adictosaltrabajo.com/2019/01/17/introduccion-al-machine-learning-con-weka/>
- Cabezas-García, E., Langer-Glass, A., Alvarez-Vázquez, L., & Bustamante, P. (1998). Perfil Sociodemográfico del aborto inducido. *Salud Pública De México*, 40(3), 265-271. Disponible en: <https://saludpublica.mx/index.php/spm/articloe/view/6085>
- Castañeda Abascal, I, & Molina Estévez, M. (1999). Factores biosociales que influyen en la aparición del aborto provocado. *Revista Cubana de Obstetricia y Ginecología*, 25(1), 55-60. Recuperado el 27 de marzo de 2022, de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0138600X1999000100011&lng=es&tlng=es
- Díaz Sánchez, V. (2003). El embarazo de las adolescentes en México. Disponible en: <https://www.medigraphic.com/pdfs/gaceta/gm2003/gms031e.pdf>
- De la Fuente L. (2015, 25 diciembre). Perfil de la mujer que aborta: de 20 a 24 años, trabaja y tiene estudios superiores a la secundaria. Disponible en: <https://www.abc.es/sociedad/abci-perfil-mujer-abortamnoticia.html>
- Fernández Borbon, H.; Gerez Mena S.; Ramirez Perez N. y Pineda Bouzon A. (2014). Caracterización de la terminación voluntaria del embarazo en adolescentes. *Rev Cubana Obstet Ginecol.* vol.40, n.2, pp.197-205. ISSN 0138-600X. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_abstract&pid=S0138-600X2014000200007
- Carbonell JL, Varela L, Velazco A, Fernández C. (1997). The use of misoprostol for termination of early pregnancy. *Contraception* 1997;55:165-168.
- Fernández Cantón S.B, Gutiérrez Trujillo G, Viguri Uribe R. (2012, febrero) La mortalidad materna y el aborto en México, párrafo 3. Disponible en:



http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-11462012000100011

Figuroa Lara, A., Aracena Genao, B., Reyes Morales, H., & Lamadrid Figuroa, H. (2012). Factores Asociados Con La Búsqueda Del Servicio De Interrupción Legal Del Embarazo En La Ciudad De México. *Salud Pública De México*, Vol.54, No. 4. Disponible en: http://scielo.org.mx/scielo.php?script=sci_arttext&pid=S0036-36342012000400010#end

Forbes Staff. (septiembre 29, 2020). ¿En qué estados de México es legal abortar? Marzo 4, 2022, de Forbes, México. Disponible en: <https://www.forbes.com.mx/noticias-en-que-estados-es-legal-abortaren-mexico/>

Gallego J. A., Navarro L. F., & Castillo L. F. (2015). Aplicación de técnicas de minería de datos en atención primaria en salud (APS) para el análisis de riesgos en mujeres gestantes de la población manizaleña atendida por Assbasalud. *Biosalud*, 14(2), 71 - 78. <https://doi.org/10.17151/biosa.2015.14.2.7>

García Álvarez CT, Seuc Jo A. (1997). Factores relacionados con la selección entre aborto o la continuación del embarazo en las adolescentes. *Salud Reproductiva en Cuba*, Vol. II. Naciones Unidas.

González de León Aguirre, D, L. Billings, D., Gasman Zylbermann, N., & Rayas Velasco. (Diciembre, 2002). *El Aborto En México*. pp. 1-24. ISBN 1-88-2220-37-4. Disponible en: <https://ccp.ucr.ac.cr/ac/gonzalez.pdf>

González de León-Aguirre D, Billings DL, Ramírez-Sánchez R. (2008). *El Aborto Y La Educación Médica En México*. *Salud Pública. Mex.*; Vol. 50, No. 3. Disponible en: <https://www.scielosp.org/pdf/spm/2008.v50n3/258-267>

González Polanco, L., & Pérez Betancourt, Y. G. (2013). *La Minería De Datos Y Su*

Aplicación En Los Estudios De Salud. *Revista Cubana de Información en Ciencias de la Salud*, Vol.24 .No.4. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2307-21132013000400010

Harden A, Ogden J. (1999). Young women's experiences of arranging and having abortions. *Sociol Health Illn*;21(4):426-444.

Islas de González, O. (2008). Evolución del aborto en México. Disponible en: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S004186332008000300006#N1

Juárez F et al., (2013). *Embarazo no planeado y aborto inducido en México: causas y consecuencias*, Nueva York: Guttmacher Institute. ISBN: 978-1-934387-12-2. Disponible en: https://www.guttmacher.org/sites/default/files/report_pdf/embarazo-no-deseadomexico0.pdf

Kuri Morales, P. (2020). *Panorama nacional del embarazo precoz en México: lecciones aprendidas en un sexenio*. Disponible en: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0016-3813202000020015

La Minería de datos Data Mining. (2018, 7 mayo). MicroStrategy. Disponible en: <https://www3.microstrategy.com/es/resource/s/introductory-guides/data-mining-explained>

Langer Glass, A., Bustamante, P., Cabezas García, E., & Álvarez Vázquez, L. (1998, 1 mayo). *Perfil Sociodemográfico del aborto*. *Salud Pública De México*. Disponible en: <https://saludpublica.mx/index.php/spm/articloe/view/6085>

M, J. (s. f.). *Minería de Datos: Análisis de clustering*. *Análisis De Datos*. Disponible en: <https://analisisdedatos.net/mineria/tecnicas/clustering/intro.php>

Machado, J., Pereiraa, E., Portela, F., & Santos, M. F. (2014, 1 diciembre).



Managing Voluntary Interruption of Pregnancy using Data Mining. ScienceDirect. Disponible en: https://www.researchgate.net/publication/268807381_Managing_Voluntary_Interruption_of_Pregnancy_Using_Data_Mining

Marbán, M.S. (2018). A Data Mining & Knowledge Discovery Process Model. Disponible en: http://cdn.intechopen.com/pdfs/5937/InTech_A_data_mining_amp_knowledge_discovery_process_model.pdf

MedlinePLus. (Julio 9, 2020). Aborto. Marzo 4, 2022, de Información de la Salud para la Salud. Disponible en: <https://medlineplus.gov/spanish/abortion.html> 6. Belinchón

Monjas, Y. (2004). Minería De Datos. P.1–4. Disponible en: <https://www.it.uc3m.es/jvillena/irc/practicas/10-11/15mem.pdf>

Montero Puñales. E, Plasencia Salgueiro A. (2016). Metodología CRISP- DM aplicada al proceso de minería de textos para agrupar documentos web. P. 2-3. Disponible en: <http://www.congresoinfo.cu/index.php/info/2016/paper/viewFile/250/220>

Navarro Robles, A. (2019). Interrupción Legal del Embarazo, experiencia en la Ciudad de México. Boletín CONAMED, Volumen 5, No. 25, Pág. 1-10. Disponible en: <http://www.conamed.gob.mx/gobmx/boletin/pdf/boletin25/b25-8.pdf>

Oracle Data Visualization. Oracle. Disponible en: <https://www.oracle.com/mx/business-analytics/data-visualization.html>

Pérez-Santos R. (1998). Algunos factores que inciden en la interrupción tardía de la gestación. Rev Cubana Obst Ginecol 14(4):64-70. 12.

Prats, J. (2005, 21 febrero). El 40% de las mujeres que abortan tienen menos de 24 años. EL PAÍS. Disponible en: https://elpais.com/diario/2005/02/21/cvalencia/1109017087_850215.html

Rodríguez Montequín, M. T., Álvarez Cabal, J. V., Mesa Fernández, J. M., & González Valdés, A. (s. f.). Metodologías Para La Realización De Proyectos De Data Mining. Recuperado 28 de marzo de 2022, de: https://www.aepro.com/files/congresos/2003pamplona/ciip03_0257_0265.2134.pdf

Rueda, J. F. V. (2019, 4 noviembre). CRISP-DM: Una Metodología Para Minería De Datos En Salud. healthdataminer.com. Disponible en: <https://healthdataminer.com/data-mining/crisp-dm-una-metodologia-para-mineria-de-datos-en-salud/>

Sánchez Vidal, K. J. (2005, 1 febrero). Nivel de participación de las personas en la decisión del aborto-Hospital Nacional Sergio Bernales. CyberTesis. Disponible en: <https://hdl.handle.net/20.500.12672/2978>

Sebastiani, M. (2018, 9 enero). El aborto como un bien social. no.43. ISSN 1886-5887 SciELO. Disponible en: http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1886-58872018000200004

Senderowicz, L., Sanhueza, P., & Langer, A. (2019, 11 junio). Escolaridad, lugar de residencia y utilización de servicios de aborto. Guttmacher Institute. Disponible en: <https://www.guttmacher.org/es/journals/ipsrh/2018/12/escolaridad-lugar-de-residencia-y-utilizacion-de-servicios-de-aborto-legal-en>

Sistema de Información de Interrupción Legal del Embarazo. (2021). Interrupción Legal del Embarazo (ILE) Estadísticas Abril 2007 – 30 de Septiembre 2021. Disponible en: <http://ile.salud.cdmx.gob.mx/wpcontent/uploads/WEB-21.pdf>



Sridharan, M. (2020, 18 abril). CRISP-DM: A framework for Data Mining & Analysis. Think Insights. Disponible en: <https://thinkinsights.net/digital/crisp-dm/>

Velasco Bozal A. Rodríguez Cárdenas I A; Rodríguez Tahuilell V; Salas Castillo III L; Díaz Curbelo A. (2008). Eficacia del Misoprostol en la interrupción del embarazo del 2do. Trimester. Rev Cubana Obstet Ginecol v.34 n.3 Ciudad de la Habana sep.-dic. 2008.