

**ANÁLISIS DE LA MORTALIDAD POR CÁNCER EN MUNICIPIOS DEL NORTE
DEL VALLE DEL CAUCA APLICANDO TÉCNICAS DE MINERÍA DE DATOS**

**STEFFANIA LÓPEZ LONDOÑO
DAVID FERNANDO AGUILAR RUIZ**



**Zarzal
Universidad Del Valle Sede Zarzal
Facultad de Ingeniería
Ingeniería Industrial
2019**

ANÁLISIS DE LA MORTALIDAD POR CÁNCER EN MUNICIPIOS DEL NORTE DEL VALLE DEL CAUCA APLICANDO TÉCNICAS DE MINERÍA DE DATOS

Proyecto de grado para optar por el título de Ingeniero Industrial

**STEFFANIA LÓPEZ LONDOÑO
DAVID FERNANDO AGUILAR RUIZ**

**DIRECTOR:
ANDRÉS GARZÓN VALENCIA**

**CODIRECTOR:
JULIAN GONZÁLEZ VELASCO**



**Zarzal
Universidad Del Valle Sede Zarzal
Facultad de Ingeniería
Ingeniería Industrial
2019**

AGRADECIMIENTOS

Doy gracias principalmente a Dios por permitirme vivir momentos y logros como este.

Agradezco infinitamente a mi familia por el apoyo constante en cada proyecto que decido realizar, por inculcar en mí los valores suficientes para ser lo que hoy en día soy, sin ellos no existiría la motivación y voluntad suficiente para cumplir cada una de las metas que me propongo.

Finalmente, gracias a todas aquellas personas que, de una forma u otra, hacen parte de mi vida y están siempre allí para impulsarme a ser mejor, con ellos este camino se ha hecho más fácil y llevadero.

A todos infinitas gracias.

Con todo el amor y el cariño, Steffania López Londoño.

Agradezco a Dios por hacer de cada día un instante de tiempo lleno de salud y vida, por brindarme la sabiduría necesaria para lograr cada una de mis metas y por permitirme tener una familia, la cual es mi principal motivación en mis proyectos personales y profesionales.

A mis padres Hernando Aguilar Rojas y Blanca Oliva Ruiz Abadía, en los que sin importar las circunstancias encuentro un apoyo incondicional, gracias por siempre creer en mí y por luchar cada día por un interés en común, el cual se centra en velar por mi bienestar.

A Maria Camila Aristizábal Cataño, en ella encuentro un gran apoyo, gracias por evitar que pierda la fe en momentos difíciles, gracias por convertir los días de trabajo en momentos de esperanza, motivación y confianza.

A Milton Alexander Londoño Román, es una persona que brinda lo mejor de sí mismo y siempre está dispuesto a ayudar a los demás; es una gran persona.

A los docentes, quienes apoyaron y dieron su mejor esfuerzo para que este proyecto se llevará a cabo de la mejor manera.

David Fernando Aguilar Ruiz

CONTENIDO

INTRODUCCION	6
1. SELECCIONAR UNA HERRAMIENTA DE MINERÍA DE DATOS QUE CONTRIBUYA AL ANÁLISIS DE LA MORTALIDAD POR CÁNCER EN MUNICIPIOS DEL NORTE DEL VALLE DEL CAUCA.....	15
1.1. TÉCNICAS DE MINERÍA DE DATOS.....	15
1.2. CONDICIONES PROPUESTAS PARA LA ELECCIÓN DE LA HERRAMIENTA DE MINERÍA DE DATOS	15
1.3. ELECCIÓN DE LA HERRAMIENTA DE ACUERDO A LA LITERATURA	15
1.3.1. WEKA.....	18
1.3.2. ORANGE.....	19
1.3.3. RAPIDMINER.....	19
1.3.4. JHEPWORK.....	19
1.3.5. KNIME.....	19
2. EJECUTAR UN DIAGNÓSTICO REFERENTE A LAS PRINCIPALES CAUSAS DE MORTALIDAD POR CÁNCER EN MUNICIPIOS DEL NORTE DEL VALLE DEL CAUCA	20
2.1. RECOLECCIÓN DE LOS DATOS	20
2.2. PREPARACIÓN DE LOS DATOS (PRE-PROCESO)	21
2.3. USO DE LA HERRAMIENTA WEKA	21
2.4. APLICACIÓN DE LOS ALGORITMOS DE CLASIFICACIÓN EN WEKA .	23
2.5. TIPOS DE CÁNCER CON MAYOR INCIDENCIA Y PRINCIPALES CAUSAS DIRECTAR DE MUERTE	26
2.6. FACTORES QUE INCIDEN EN EL RIESGO DE CONTRAER CÁNCER.	29
3. ANALIZAR LOS RESULTADOS CON BASE EN LA IDENTIFICACIÓN DE PATRONES O COMPORTAMIENTOS DE DATOS DE INTERÉS.....	31
3.1. ESTRUCTURA DEL ÁRBOL DE DECISIONES	31
3.1.1. Raíz.....	31
3.1.2. Nodos internos.....	31
3.1.3. Hojas	31
3.2. MATRIZ DE CONFUSIÓN	31
3.3. APLICACIÓN PARA EL USO DE RESULTADOS	33
3.3.1. Atributos que contiene la aplicación	33
3.3.2. Funcionamiento del aplicativo web.....	34

3.3.3.	Árbol de decisión en la programación	35
3.3.4.	Resultado de la predicción	35
4.	CONCLUSIONES	36
5.	FUTUROS TRABAJOS.....	38
6.	BIBLIOGRAFÍA.....	39
7.	ANEXOS.....	42
7.1.	Anexo 1. ÁRBOL DE DECISIONES	42
7.2.	Anexo 2. MATRIZ DE CONFUSIÓN	47

TABLAS

Tabla 1 Factores que Inciden en el Riesgo de Contraer Cáncer	8
Tabla 2 Tipos de Cáncer	9
Tabla 3 Estadísticas del Cáncer	11
Tabla 4 Resumen de Artículos	16
Tabla 5 Evaluación de Algoritmos	25
Tabla 6 Tipos de Cáncer con Mayor Incidencia	26
Tabla 7 Principales Causas de Muerte	27
Tabla 8 Factores que inciden en el riesgo de contraer cáncer y que se pueden evitar	29
Tabla 9 Factores que inciden en el riesgo de contraer cáncer y que no se pueden evitar	30
Tabla 10 Instancias incorrectamente clasificadas	33

ILUSTRACIONES

Ilustración 1 Mortalidad por neoplasias en hombres y mujeres	10
Ilustración 2 Elementos a Evaluar.....	17
Ilustración 3 Cumplimiento de los Elementos a Evaluar	17
Ilustración 4 Funcionalidades de las Herramientas.....	18
Ilustración 5 Aplicación del Filtro SubSetByExpression	22
Ilustración 6 Selección de Atributos que Influyen en la Causa Directa de Muerte .	22
Ilustración 7 Evaluación del Algoritmo y Selección de la Etiqueta	23
Ilustración 8 Ejemplo de Evaluación de Algoritmos	24
Ilustración 9 Resultados de la Clasificación	25
Ilustración 10 Ejemplo, Matriz de confusión.....	32
Ilustración 11 Interfaz de formulario de predicción.....	34
Ilustración 12 Ejemplo, programación relacionada con el árbol de decisión.....	35
Ilustración 13 Resultado de la predicción	35

INTRODUCCION

La Organización Mundial de la Salud afirma que “El cáncer es un proceso de crecimiento y diseminación incontrolado de células, que puede aparecer prácticamente en cualquier lugar del cuerpo” (*Organización Mundial de la Salud - OMS, 2019*). Así mismo, asegura que dicha enfermedad es la segunda causa de muerte en el mundo y cerca del 70% de ellas se registran en países de ingresos medios y bajos, de los cuales, solo uno de cada cinco dispone de datos necesarios para impulsar políticas de lucha contra la enfermedad mencionada. (*Organización Mundial de la Salud - OMS, 2018*)

Por su parte, el Instituto Nacional del Cáncer en un comunicado realizado, dice que “En general, no es posible saber con exactitud por qué una persona padece cáncer y otra no. Pero la investigación ha indicado que ciertos factores de riesgo pueden aumentar la posibilidad de una persona de padecer cáncer”. Dentro de la lista de factores que inciden en el riesgo de contraer cáncer, algunos de ellos se pueden evitar mientras que otros no (*Instituto Nacional del Cáncer, 2015*), tales factores se muestran en la *Tabla 1*.

Tabla 1 Factores que Inciden en el Riesgo de Contraer Cáncer

FACTORES QUE SE PUEDEN EVITAR	FACTORES QUE NO SE PUEDEN EVITAR
Hormonas	Sustancias químicas en el ambiente
Radiación	Gérmenes infecciosos
Luz solar	Inflamación crónica
Obesidad	Inmunosupresión
Alcohol	Tabaco
Dieta	Edad

Fuente: Autores con base en información suministrada por el Instituto Nacional del Cáncer

Tras identificar algunos de los factores que inciden en el riesgo de contraer esta enfermedad, la American Cancer Society nombra los diferentes tipos de cáncer que se pueden llegar a presentar debido a la exposición a dichos factores; los cuales son expuestos en la *Tabla 2*. (*American Cancer Society, 2019*)

Tabla 2 Tipos de Cáncer

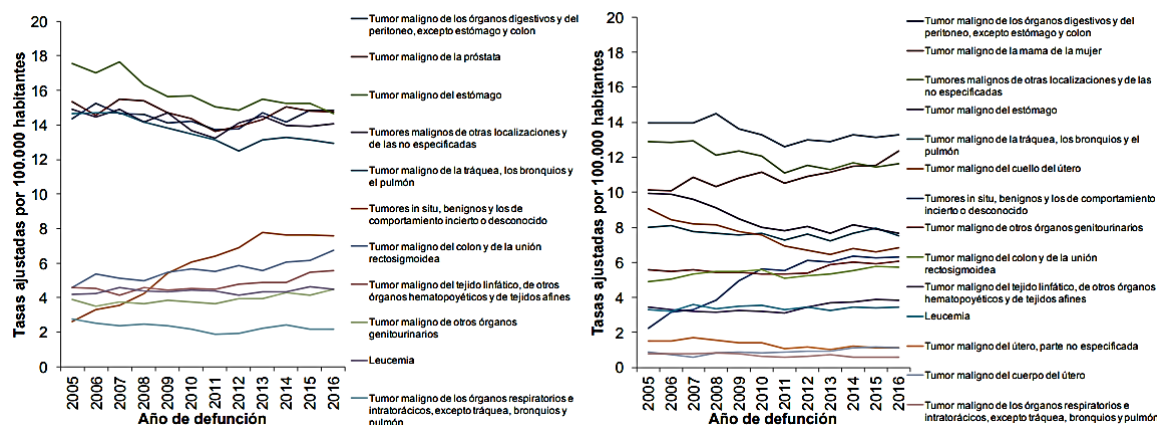
PRINCIPALES TIPOS DE CÁNCER	
Sarcoma en tejidos blandos	Tumores de encéfalo y de medula espinal
Cáncer de vesícula biliar	Cáncer de cuello uterino o cervicouterino
Cáncer de colon o recto	Cáncer de laringe y de hipofaringe
Cáncer de endometrio	Cáncer de orofaringe y de cavidad oral
Cáncer de estómago	Cáncer de pulmón
Cáncer de esófago	Cáncer de páncreas
Cáncer de hueso	Cáncer de riñón
Cáncer de hígado	Cáncer de seno
Cáncer de ovario	Cáncer de piel
Cáncer de pene	Sarcoma uterino
Cáncer de próstata	Cáncer de testículo
Mieloma múltiple	Cáncer de tiroides
Cáncer de ojo	Cáncer de vagina
Leucemia	Cáncer de vejiga
Linfoma	Cáncer de vulva

Fuente: Autores con base en la información suministrada por La American Cancer Society

Dentro de este marco, en Colombia, según el Ministerio de Salud y Protección Social y el Instituto Nacional de Cancerología, entre los tipos de cáncer más frecuentes según las cifras de mortalidad para el año 2010 se encuentra el cáncer de mama como la primera causa de muerte (2.381), seguido por el cáncer cervicouterino (1.892), estómago (1.709), pulmón (1.606), colon y recto (1.456), esto para el caso de las mujeres. Entre los hombres, la mortalidad por cáncer para este mismo año, fue encabezada por los tumores malignos del estómago (2.796), seguido por los de próstata (2.431), pulmón (2.357), colon y recto (1.261) y las leucemias (890). El cáncer también se presenta entre los niños y jóvenes representando la segunda causa de muerte entre la población de 1 a 15 años de edad, registrándose alrededor de 500 muertes por cáncer en infantes. (*Ministerio de Salud y Protección Social, 2019*). De acuerdo al (ASIS) a nivel nacional, las neoplasias son la tercera causa directa de muerte siendo los tumores malignos de los órganos digestivos y del peritoneo la primera causa de mortalidad en hombres y mujeres, seguido por el tumor maligno de estómago y el cáncer de próstata. (Dirección de Epidemiología y Demografía, 2018).

En la *ilustración 1* se muestra la Mortalidad por neoplasias en hombres (izquierda) y mujeres (derecha), 2005-2016.

Ilustración 1 Mortalidad por neoplasias en hombres y mujeres



Fuente: ASIS 2018

Así mismo, en el Valle del Cauca de acuerdo a las cifras entregadas en el último informe del Análisis de Situación de Salud (ASIS) del año 2017, las neoplasias ocupan el tercer lugar como causa de muerte en la población, alcanzando su máximo en el año 2005 con 111,41 muertes por cada 100.000 habitantes, y cerrando el informe con 105,15 muertes en el 2015. (*Secretaría Departamental De Salud del Valle del Cauca, 2018*). Según información encontrada en la página web *ENCOLOMBIA.COM* y en relación con los factores de riesgo; en el departamento del Valle del Cauca el sector económico predominante es el agrícola, este cuenta con una variedad de cultivos como caña de azúcar, café, algodón, soya y sorgo; “la principal industria es la azucarera, junto con la producción de papel, químicos, maquinaria y equipos; siendo Yumbo, Santiago de Cali y Palmira las principales sedes industriales. Otro sector de gran desarrollo en los últimos tiempos es la construcción, la explotación de la ganadería bovina, porcina, la avicultura y la pesca. Por otro lado, los recursos mineros están representados por carbón, caliza, mármoles, oro, plata, platino, hierro, asbesto, entre otros” (*Departamento del Valle del Cauca, 2019*); lo anterior, crea diferentes factores que impactan de manera negativa en la salud de las personas que viven en el departamento y en sus alrededores, aumentando el riesgo de contraer cáncer debido a las diversas actividades que se ejecutan en los diferentes sectores económicos, como son:

- *Aplicaciones de insumos a los diferentes cultivos.* Estas aplicaciones contaminan el aire y agua, el cual es inhalado o consumido por las personas.
- *Actividades industriales.* Para realizar la limpieza al final de las actividades empresariales, en muchas ocasiones se utilizan químicos o sustancias tóxicas como ácidos, metales pesados, sales y restos de materiales en descomposición, al limpiar el área, muchas veces no se les da un adecuado tratamiento a las aguas residuales, que son vertidas directamente a los ríos, lagos o mares (*Torey, 2019*); finalmente el líquido contaminado puede ser

utilizado para usos del hogar o para regar cultivos que después de ser cosechados, llegarán a las personas para su posterior consumo.

- *Exposición a los recursos mineros.* Personas que manipulan carbón, asbestos, hierro, entre otras.

A nivel municipal y con base en el último reporte del ASIS publicado en el año 2018, para Zarzal Valle del Cauca y municipios aledaños, el cáncer representa una de las principales causas de muerte. En la *Tabla 3* se muestra la posición que tiene el cáncer con respecto a las demás causas de muerte (*Ranking*), el número de muertes por cada 100.000 habitantes en el último año (*Número de muertes por cáncer*) y el cáncer con mayor influencia en cada municipio (*Tipo de cáncer con mayor tasa*). (Secretaría Departamental De Salud del Valle del Cauca, 2018)

Tabla 3 Estadísticas del Cáncer

Municipio	Ranking	# de muertes por cáncer	Tipo de cáncer con mayor Tasa
Bolívar	Causa #3	118,5	Tumores in situ, benignos y los de comportamiento incierto o desconocido.
Roldanillo	Causa #3	142,41	Tumor maligno de la próstata
La Unión	Causa #3	115,01	Tumor maligno de la próstata
La Victoria	Causa #4	60,82	Tumor maligno de los órganos digestivos y del peritoneo, excepto estómago y colon
Toro	Causa #4	68	Tumor maligno de los órganos digestivos y del peritoneo, excepto estómago y colon
Zarzal	Causa #3	75,74	Tumores in situ, benignos y los de comportamiento incierto o desconocido

Fuente: Autores con base en la información suministrada por el ASIS 2017

Con el fin de disminuir el impacto de la enfermedad, la república de Colombia junto con el Ministerio de Salud y Protección Social y el Instituto Nacional de Cancerología, crean un plan nacional para el control del cáncer entre el 2012-2020, mediante la reducción del riesgo a padecerlo, la disminución de la tasa de mortalidad y la mejora en la calidad de vida para aquellas personas que lo padecen. (Ministerio de Salud y Protección Social; Instituto Nacional de Cancerología, 2012)

Si bien es cierto, el área de la salud es uno de los sectores en donde a diario se crean grandes cantidades de datos, pues podemos ver como las personas acuden a los centros de salud para recibir todo tipo de atención, ya sea por una urgencia, una cita médica o una toma de exámenes; es así como se van agregando datos a sus sistemas de información y se quedan en la *“historia clínica”* de cada habitante; dichos datos, en su mayoría no son analizados y por lo tanto no se saca el máximo

provecho para encontrar alguna aplicabilidad a la información oculta, que además puede servir para solucionar problemas o mejorar situaciones alrededor del mundo.

Lo anterior, crea la necesidad de analizar grandes cantidades de datos, que en este caso son generados en el sector salud; un campo de la estadística y las ciencias de la computación que hace posible dicho análisis y que además permite convertir los datos en información importante tanto en el presente como en el futuro, es la minería de datos. Las técnicas de la minería de datos tienen como objetivo descubrir patrones, perfiles y tendencias a través del análisis de los datos utilizando tecnologías de reconocimiento de patrones, redes neuronales, lógica difusa, algoritmos genéticos, y otras técnicas avanzadas de análisis de datos. (López, 2007). De acuerdo al libro escrito por Perez & Santín (2008) titulado “Minería de Datos: conceptos, técnicas y sistemas”, las técnicas de minería de datos se clasifican inicialmente en técnicas predictivas y descriptivas, en “las técnicas predictivas especifican el modelo para los datos con base en un conocimiento teórico previo”, ya que el modelo es entrenado previamente para predecir la variable, mientras que, “en las técnicas descriptivas, no se asigna ningún papel predeterminado a las variables”, es decir, son grupos de datos que no son conocidos con anterioridad. (Perez & Santín, 2008)

Al conocer el gran impacto que está generando la minería de datos dentro de cualquier área de estudio y al ver el problema de mortalidad por cáncer, se decide realizar este proyecto con el fin de brindar un análisis, mediante el descubrimiento de patrones, que permita conocer los principales tipos de cáncer que conllevan a las altas tasas de mortalidad causadas por dicha enfermedad.

Basados en la literatura, con respecto a la problemática considerada, los autores Dursun, Glenn, & Amit, (2005), en su trabajo de investigación “*Predicting breast cancer survivability: a comparison of three data mining methods*”, utilizan dos algoritmos populares de minería de datos, redes neuronales artificiales y árboles de decisión, esto con el fin de desarrollar modelos de predicción utilizando un gran conjunto de datos (más de 200.000 casos), sus resultados indicaron que el árbol de decisión es el mejor predictor con un 93.6% de precisión en la muestra retenida. Por otro lado, Soni, Ansari, Sharma, & Soni, (2011), en su investigación “*Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction*”, utilizan la encuesta de revistas y publicaciones en los campos de la medicina, la informática y la ingeniería. Dicha investigación se centró en las publicaciones más recientes; tres clasificadores o algoritmos se usaron para diagnosticar la presencia de enfermedad cardíaca en pacientes (*Decision Tree, Naïve Bayes y Classification via clustering*). Como resultado obtuvieron que el Árbol de decisión supera a las otras dos técnicas de minería de datos.

También, Timarán & Yépez (2016), en su artículo “*Caracterización de la supervivencia de mujeres con cáncer invasivo de cuello uterino usando minería de datos*”, presentan el proceso de descubrimiento de patrones de supervivencia en

casos de mujeres diagnosticadas con cáncer invasivo de cuello uterino en el periodo 1998-2003 los cuales fueron observados hasta el año 2007. Con base en los datos almacenados en el Registro Poblacional de Cáncer del municipio de Pasto; mediante la técnica de árboles de decisión, obtuvieron que el patrón principal descubierto es aquel que caracteriza a una mujer con cáncer invasivo de cuello uterino como sobreviviente, si sobrepasa los 52 meses después del momento del diagnóstico del cáncer.

Por último, *Calleja (2010)* en su trabajo de pregrado “*Minería de datos con Weka para la predicción del precio de automóviles de segunda mano (2010)*” de la Universidad Politécnica de Valencia, con la finalidad de estudiar y profundizar en la Minería de Datos, desarrolla su proyecto en dos partes la primera consta de un estudio de los conceptos más importantes de la metodología, sus orígenes y el futuro de las técnicas de tratamiento de datos, la segunda parte consistió en la aplicación de dichas técnicas en un caso de estudio real mediante los árboles de decisión para extracción de información de los datos obtenidos a través de una aplicación en Microsoft Access, *Calleja* deja abierta la posibilidad de que en un futuro, puedan acceder todas las personas interesadas en hacer uso de la aplicación, y para ello propone cambiar de plataforma y realizarla como aplicación web, “La aplicación en Access, podría trasladarse a cualquier otra plataforma o entorno de programación para facilitar su utilización. Realizarla como aplicación web sería idóneo”.

Este proyecto se realizó bajo un método de investigación de tipo exploratorio, en la región, el cual no ha sido estudiado a detalle. Según la página web *Universia* “la investigación de tipo exploratoria se realiza para conocer el tema que se abordará, lo que nos permite familiarizarnos con algo que hasta el momento desconocíamos” (*Universia Costa Rica, 2017*).

Seguidamente se plantearon tres objetivos que permitieron llegar a la solución de un objetivo principal que consiste en “Proponer un estudio de la mortalidad por cáncer en municipios del norte del Valle del Cauca, aplicando técnicas de minería de datos que faciliten el análisis de la información para la extracción de patrones de comportamiento y tendencias”. Como objetivos específicos se tiene

1. “Seleccionar una herramienta de Minería de Datos que contribuya al análisis de la mortalidad por cáncer en municipios del norte del Valle del Cauca.” En esta etapa del proyecto, se investigaron las diferentes herramientas de minería de datos con el fin de encontrar la que más se adapte al objetivo que se quiere alcanzar, esto teniendo en cuenta los diferentes trabajos encontrados en la literatura.
2. “Ejecutar un diagnóstico referente a las principales causas de mortalidad por cáncer en municipios del norte del Valle del Cauca.” Se determinaron cuáles son los tipos de cáncer con más influencia dentro de las bases de datos de

mortalidad proporcionadas, las cuales fueron obtenidas en municipios del norte del Valle del Cauca; para ello se tomaron los resultados del objetivo anterior y se procedió a ejecutar el diagnóstico para ser clasificados de mayor a menor incidencia los tipos de cáncer más frecuentes en dicha población y las causas directas de muerte en cada uno de ellos.

3. “Analizar los resultados con base en la identificación de patrones o comportamientos de datos de interés.” Teniendo en cuenta que es una investigación de tipo exploratorio, en esta etapa se analizaron los resultados del estudio para determinar qué tipo de aplicabilidad se le puede dar con respecto a la técnica seleccionada en el objetivo número uno.

1. SELECCIONAR UNA HERRAMIENTA DE MINERÍA DE DATOS QUE CONTRIBUYA AL ANÁLISIS DE LA MORTALIDAD POR CÁNCER EN MUNICIPIOS DEL NORTE DEL VALLE DEL CAUCA

En esta etapa del proyecto se investigaron diferentes técnicas y herramientas de minería de datos con el fin de encontrar la que más se adapte al objetivo que se quiere alcanzar, esto teniendo en cuenta los diferentes trabajos encontrados en la literatura.

1.1. TÉCNICAS DE MINERÍA DE DATOS

Según el libro publicado por *Perez & Santín (2008)*, las técnicas de minería de datos están enfocadas al descubrimiento de patrones dentro de grandes cantidades de datos, dichas técnicas son las predictivas y las descriptivas.

En las técnicas predictivas el modelo se crea de acuerdo a un conocimiento teórico previo, es decir, antes de arrojar cualquier resultado los datos primero son analizados y reconocidos, y sobre dicho aprendizaje se da la solución o la creación del modelo. Dentro de estas técnicas se encuentran los árboles de decisión, algoritmos genéticos, regresión, métodos Bayesianos y las redes neuronales. (*Perez & Santín, 2008*)

Por otro lado, en las técnicas descriptivas los modelos se crean automáticamente, es decir, no existe un reconocimiento previo de los datos. En las técnicas descriptivas se encuentra el clustering y segmentación y las reglas de asociación y dependencia. (*Perez & Santín, 2008*)

1.2. CONDICIONES PROPUESTAS PARA LA ELECCIÓN DE LA HERRAMIENTA DE MINERÍA DE DATOS

Para elegir la herramienta, se establecen características como que el software sea de libre uso, capacidad para trabajar con la cantidad de datos que se tenga, la facilidad de instalación, la facilidad para configurar y la interacción con el usuario.

1.3. ELECCIÓN DE LA HERRAMIENTA DE ACUERDO A LA LITERATURA

A continuación, en la *Tabla 4* se presenta un resumen de los artículos más relevantes que fueron consultados a lo largo de este proyecto, esta contiene el artículo de donde se obtuvo la información, el problema que fue abordado, la herramienta de minería de datos que fue usada, el tipo de técnica (*predictiva/descriptiva*) y el nombre de esta.

Tabla 4 Resumen de Artículos

Artículo (Cita)	Problema Abordado	Herramienta	Tipo de Técnica	Nombre de la Técnica
Dursun, Glenn, & Amit (2005)	Predicción de la supervivencia del cáncer de mama	WEKA	Predictiva	Árbol de decisión
Oztekin, Delen, & Kong (2009)	Predicción de la supervivencia del injerto para pacientes con trasplante de corazón y pulmón	X	Predictiva	Árboles de decisión – Regresión logística – Redes neuronales
Soni, Ansari, Sharma, & Soni (2011)	Predicción de enfermedades del corazón	WEKA	Predictiva	Árboles de decisión
Timarán & Yépez, (2016)	Descubrimiento de patrones de supervivencia	WEKA	Predictiva	Árboles de decisión
Calleja, (2010)	Predicción del precio de automóviles	WEKA	Predictiva	Árboles de decisión
Vega et al., (2019)	Predicción y tratamiento de enfermedades	WEKA - ORANGE	Predictiva	Árboles de decisión

Fuente: Autores

De acuerdo a los artículos antes expuestos, se encontró uno que sería de gran ayuda para determinar la herramienta apropiada con base en las condiciones antes propuestas. Dicho artículo propone una metodología de aplicación de algoritmos de predicción, utilizando técnicas de minería de datos, allí muestran un importante análisis para elegir la herramienta apropiada; comparando cinco tipos de software de minería de datos de acuerdo a catorce factores de evaluación. (Vega, Rosano, López, Cendejas, & Ferreira, 2019)

A continuación, se presenta la *Ilustración 2*, en donde se encuentran los elementos a evaluar, posteriormente en la *Ilustración 3*, se muestra el cumplimiento de estos elementos por cada uno de los softwares de minería de datos contemplados en dicho artículo.

Ilustración 2 Elementos a Evaluar

ELEMENTOS A EVALUAR			
1	Flexibilidad	8	Fácil de configurar
2	Portabilidad	9	Amigable
3	Interoperabilidad	10	Íconos con ayuda
4	Manual Técnico	11	Seguridad
5	Manual de Usuario	12	Actualizaciones
6	Ayuda en Línea	13	Soporte Técnico
7	Fácil de instalar	14	Independencia de Hw

Fuente: Vega et al., (2019)

Ilustración 3 Cumplimiento de los Elementos a Evaluar

Software de Minería de Datos	1	2	3	4	5	6	7	8	9	10	11	12	13	14
WEKA	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
ORANGE	✓		✓	✓	✓	✓		✓		✓	✓	✓	✓	✓
RAPIDMINER	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
JHEPWORK	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓			✓
KNIME	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓		✓	✓

Fuente: Vega et al., (2019)

Vega et al., (2019), hacen un estudio adicional para comparar las funcionalidades de las cinco herramientas, en el cual se tienen en cuenta el procesamiento de datos, si cuenta con métodos predictivos, métodos de descripción de datos, graficación y validación del modelo. En la *Ilustración 4*, se muestran los resultados obtenidos.

Ilustración 4 Funcionalidades de las Herramientas

Herramientas de Minería de Datos	Procesamiento de datos	Modelos Predictivos	Método de descripción de datos	Graficación	Validación del modelo
Weka	✓	✓	✓	✓	✓
Orange	✓	✓	✓	✓	✓
RapidMiner	✓	✓		✓	
JHepWork	✓	✓		✓	
Knime	✓	✓		✓	

Fuente: Vega et al., (2019)

Finalmente utilizan dos softwares, Weka y Orange; ya que estos dos son los que cumplen el mayor número de los *elementos a evaluar y funcionalidades de la herramienta*, además utilizaron árboles de decisión con el fin de generar resultados del estudio de manera gráfica.

A continuación, se da una breve descripción de los softwares expuestos anteriormente.

1.3.1. WEKA

Weka es un software de código abierto emitido bajo la Licencia Pública General de GNU, que contiene una colección de algoritmos de aprendizaje automático para tareas de minería de datos con herramientas para la preparación de datos, clasificación, regresión, agrupación, extracción de reglas de asociación y visualización. (Waikato, s.f.)

Según (Mark, 2009), WEKA goza de una amplia aceptación tanto en el mundo académico como en los negocios. Tiene una comunidad activa y se ha descargado más de 1.4 millones de veces desde que se colocó en Source-Forge en abril de 2000.

Las versiones estables solo reciben correcciones de errores, mientras que la versión de desarrollo recibe nuevas características. Weka 3.8 y 3.9 cuentan con un sistema de administración de paquetes que facilita a la comunidad de Weka agregar nuevas funcionalidades. El sistema de gestión de paquetes requiere una conexión a Internet para descargar e instalar paquetes. (Waikato, Descargando e instalando Weka, s.f.)

1.3.2. ORANGE

Orange, es una herramienta de código abierto para análisis y visualización de datos donde la minería de datos se hace a través de programación visual o por código de Python. Consiste en una interfaz en la que el usuario coloca widgets y crea flujos de trabajo para el análisis de datos. Estos widgets otorgan funcionalidades como leer datos, mostrar una tabla, selección de variables, comparación de algoritmos, visualización de elementos entre otros. (*DIGITAL, s.f.*)

Es una herramienta de minería de datos tanto para principiantes como para expertos en datos científicos. Gracias a su interfaz, los usuarios pueden centrarse en el análisis de datos en lugar de en la codificación laboriosa, haciendo que la construcción compleja de análisis de datos sea simple. (*Ljubljana, s.f.*)

1.3.3. RAPIDMINER

RapidMiner es una herramienta de Minería de Datos ampliamente usada y probada a nivel internacional en aplicaciones empresariales, de gobierno y academia, implementa más de 500 técnicas de pre-procesamiento de datos, modelación predictiva y descriptiva, métodos de prueba de modelos y visualización de datos.

Esta herramienta ha sido utilizada en más de cuarenta países y en compañías como Ford, Honda, E.ON, Nokia, IBM, Cisco, Hewlett Packard, Elexso, Akzo Nobel, PharmaDM, Bank of America, Merrill Lynch, entre muchas otras. (*Barra, s.f.*)

1.3.4. JHEPWORK

jHepWork (jWork) es un entorno de desarrollo Java para cálculos científicos, análisis de datos y visualización de datos. (*García, 2011*) Es un entorno para computación científica, análisis de datos y visualización de datos para los científicos, ingenieros y estudiantes. El programa es completamente multiplataforma e integrado con la lengua scripting de Jython (Python). (*Chekanov, 2013*)

1.3.5. KNIME

KNIME Analytics Platform es el software de código abierto para crear aplicaciones y servicios de ciencia de datos. KNIME hace que la comprensión de datos y el diseño de flujos de trabajo de ciencia de datos y componentes reutilizables sean accesibles para todos. Está construido sobre la plataforma Eclipse y es extremadamente flexible y potente.

Este software integra diversos componentes para aprendizaje automático y minería de datos a través de su concepto de fraccionamiento de datos. La interfaz gráfica de usuario permite el montaje fácil y rápido de nodos para pre-procesamiento de datos (ETL: extracción, transformación, carga), para el análisis de datos, modelado

y visualización. Desde 2006 es utilizado en la investigación farmacéutica, y en otras áreas como análisis de datos de cliente de CRM, inteligencia de negocio y análisis de datos financieros (*KNIME*, s.f.).

Con la información brindada por el estudio, con respecto a la elección de la mejor herramienta de minería de datos, se pudo observar que hay tres softwares que se ajustan a los requerimientos propuestos, estos son Weka, Orange y RapidMiner. Sin embargo, al revisar la tabla respecto al cumplimiento de los elementos a evaluar, Weka es la herramienta que cumple con la mayoría de los criterios establecidos, careciendo únicamente de la ayuda en línea.

Por tanto, la herramienta de Minería de Datos que se utilizó para el análisis de la tasa de mortalidad por cáncer en municipios del norte del Valle del Cauca es WEKA, además los diferentes artículos consultados convergen en la utilización de este software.

2. EJECUTAR UN DIAGNÓSTICO REFERENTE A LAS PRINCIPALES CAUSAS DE MORTALIDAD POR CÁNCER EN MUNICIPIOS DEL NORTE DEL VALLE DEL CAUCA

En este capítulo se hizo uso de la herramienta seleccionada y de la base de datos creada a partir de la información suministrada por diferentes municipios del norte del Valle del Cauca. Adicionalmente, se determinó cuáles son los tipos de cáncer con más influencia dentro de dicha base de datos de mortalidad y finalmente se ejecutó el diagnóstico para con su resultado, clasificar de mayor a menor los tipos de cáncer más frecuentes en dicha población y las causas directas de muerte en cada uno de ellos.

A continuación, se muestran las diferentes actividades que conducen al logro de este objetivo; estas actividades van desde la recolección de los datos, hasta la aplicación de los algoritmos de clasificación a la base de datos organizada.

2.1. RECOLECCIÓN DE LOS DATOS

Los datos con los cuales se desarrolla el presente proyecto fueron solicitados a diferentes hospitales de municipios del norte del Valle del Cauca a través de una carta certificada por la Universidad y dirigida a la gerencia del hospital de cada municipio, en la cual se solicitaba el permiso para acceder a sus bases de datos de mortalidad.

Los municipios a los cuales se les hizo la solicitud fueron Bolívar, Roldanillo, La Unión, La Victoria, Toro y Zarzal; de dicha solicitud se obtuvo una respuesta satisfactoria y a tiempo de los municipios de Bolívar, La Unión, La Victoria y Zarzal.

Las bases de datos de mortalidad suministradas abarcan un periodo de tiempo que va desde el año 2013 hasta el tercer trimestre del año 2018.

2.2. PREPARACIÓN DE LOS DATOS (PRE-PROCESO)

La preparación de las bases de datos obtenidas en los municipios, se realizó de la siguiente manera:

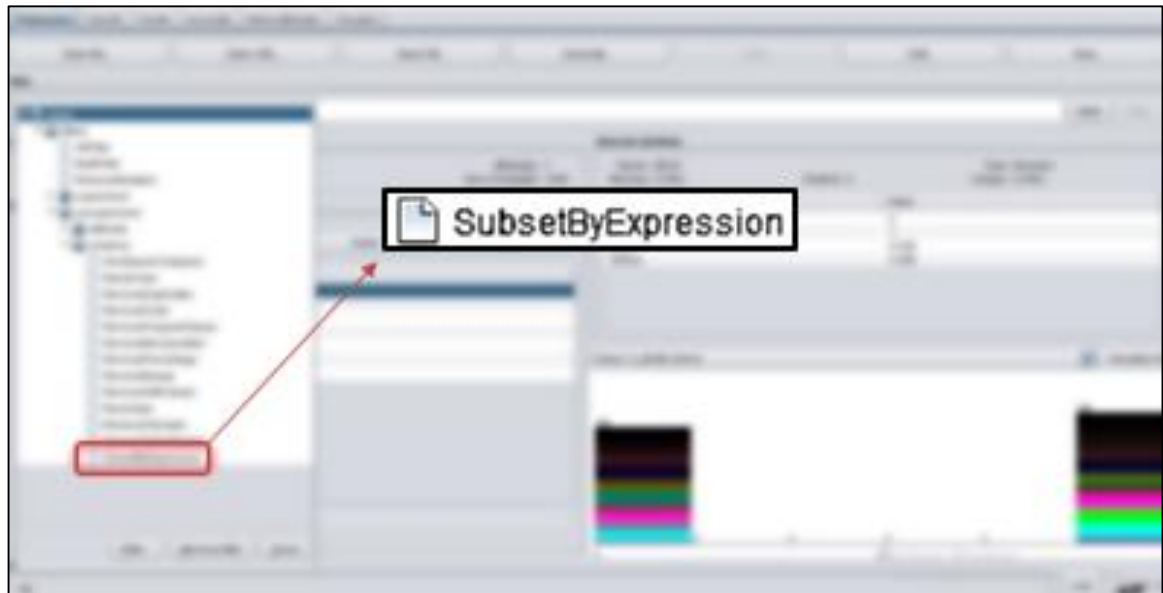
- Se depuró la información que contenía datos de pacientes que no padecieron algún tipo de cáncer durante su vida, ya que el estudio es directamente enfocado al análisis de la tasa de mortalidad por cáncer.
- De estos pacientes, se seleccionaron los atributos que podrían influir en el análisis de los datos, tales como edad, sexo, ocupación, nivel de educación, régimen de seguridad, nombre de la entidad de salud, causa directa de muerte, antecedentes del fallecido y el tipo de cáncer que padecía (*Causa base*).
- Con ello, se creó una nueva base de datos la cual contiene la información de los atributos antes seleccionados para cada paciente; cabe resaltar que dicha base de datos, contiene los pacientes de todas las bases de datos suministradas por los hospitales de cada municipio.

2.3. USO DE LA HERRAMIENTA WEKA

Con la nueva base de datos agrupada, se procedió a aplicar la herramienta seleccionada para el análisis de los datos, sin embargo, para hacer uso de WEKA, primero se llevaron a cabo los siguientes pasos:

- Debido a que la base de datos se encuentra en un archivo de Excel (.xlsx), debió convertirse a un archivo delimitado por comas (.csv), ya que este es el formato que WEKA trabaja.
- Se desplegó la nueva base de datos en el software y se aplicaron los siguientes filtros:
 - En la *Ilustración 5*, se muestra la aplicación del filtro “*SubSetByExpression*”, usado con el fin de eliminar las filas de la base de datos que contengan atributos con casillas vacías o sin información, ya que estas pueden generar errores en el momento de ejecutar la clasificación.

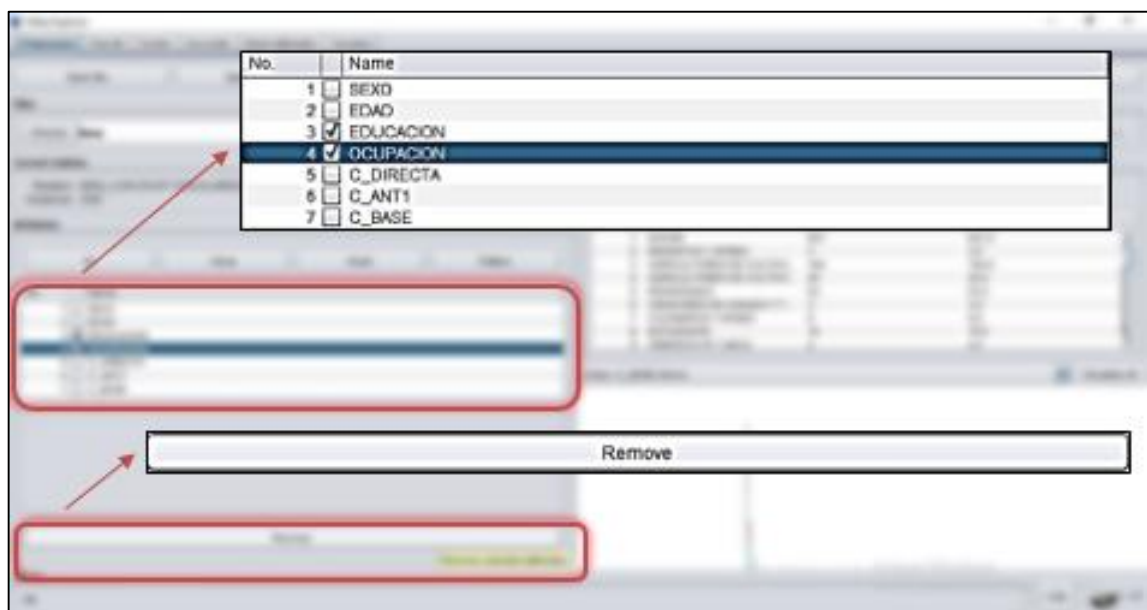
Ilustración 5 Aplicación del Filtro SubSetByExpression



Fuente: Autores

- En la *Ilustración 6*, se muestra la selección de aquellos atributos que influyen en el de “causa directa de muerte”, ya que este atributo será usado como etiqueta en el modelo. Los demás que no influyen en el atributo seleccionado, se eliminaron con la opción de Weka “Remove”.

Ilustración 6 Selección de Atributos que Influyen en la Causa Directa de Muerte



Fuente: Autores

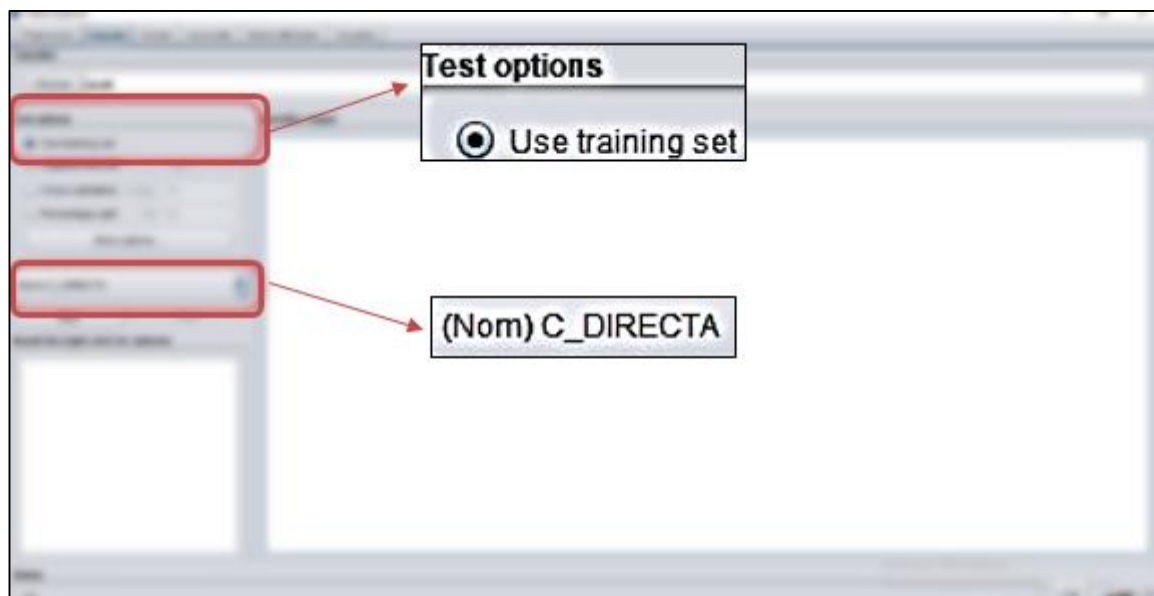
- Al observar que del atributo tipo de cáncer “(C_BASE)”, algunos cánceres tienen poca frecuencia dentro de la base de datos, se procede a filtrar aquellos que tienen una frecuencia significativa. Para esto se exporta la base de datos a Excel y se eliminan aquellos cánceres que, por su poca frecuencia, no se tomarán en cuenta para la siguiente etapa del proyecto.

2.4. APLICACIÓN DE LOS ALGORITMOS DE CLASIFICACIÓN EN WEKA

Con esta nueva base de datos, se realizó la selección del algoritmo teniendo en cuenta el tipo de datos con los que se cuenta, es decir, si en la base de datos hay atributos con valores numéricos o por el contrario tiene como valor un texto, WEKA activa automáticamente algoritmos que se ajustan a cada tipo; para esto se realiza lo siguiente:

- En la *Ilustración 7*, se observa que en el área llamada “*Test options*”, se seleccionó el evaluador “*Use training set*” y seguidamente el atributo que se tomó como etiqueta en el modelo que para este caso fue *C_DIRECTA* que hace referencia a la causa directa de muerte. El evaluador “*Use training set*”, según Valeria Guevara en su publicación en *Scientific Databases*, “*entrena el método con todos los datos disponibles y aplica el resultado sobre la misma colección de datos*”. (Guevara, 2019)

Ilustración 7 Evaluación del Algoritmo y Selección de la Etiqueta



Fuente: Autores

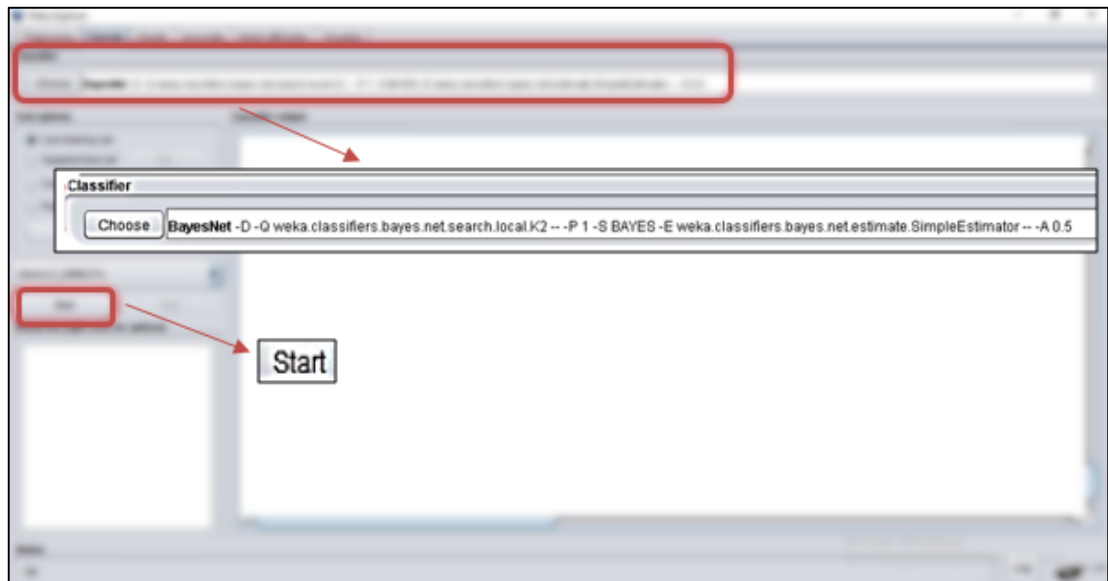
- Posteriormente, se analizaron todos los algoritmos predictivos que WEKA permite aplicar al tipo de base de datos que se tiene, de allí se compararon

los resultados y se determinó cuál es el que mejor resultado arroja con respecto al porcentaje de instancias correctamente clasificadas. En la pestaña “Classify” se selecciona la opción “Classifier”, donde se agrupan los algoritmos en diferentes clases. Para este proyecto se aplican las técnicas Bayesianas y los Árboles de decisión, ya que como se mencionó anteriormente, estos son los que se aplican en las técnicas predictivas.

A continuación, se muestra un ejemplo de lo que se realizó para aplicar los diferentes algoritmos y por último se expone una tabla comparando todos los algoritmos evaluados:

- *Seleccionar el algoritmo:* Se inició con las técnicas bayesianas, de las cuales se seleccionó el primer algoritmo a evaluar de dicha técnica, en este caso el “BayesNet”, iniciando la clasificación con el botón “Start”. En la *Ilustración 8* se muestra un ejemplo de la forma en que se evaluó dicho algoritmo junto con los algoritmos que pertenecen a la técnica de bayes y los que se evaluaron en la técnica de árboles de decisión.

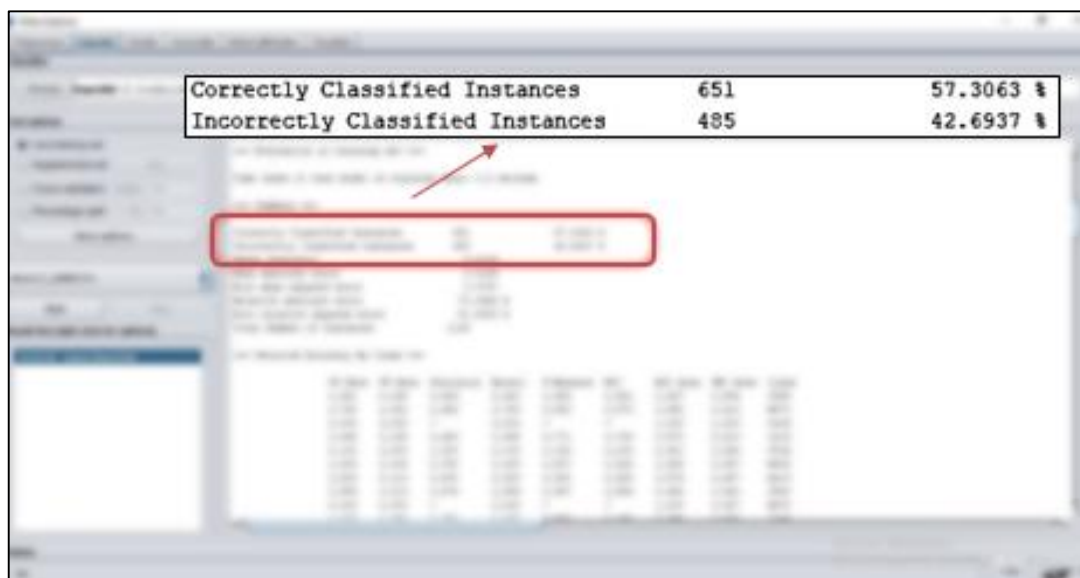
Ilustración 8 Ejemplo de Evaluación de Algoritmos



Fuente: Autores

- En el cuadro “Classifier output” como se observa en la *Ilustración 9*, se exponen los resultados de la clasificación, para este caso se tendrán en cuenta dos resultados: las instancias correctamente clasificadas (*Correctly Classified Instances*) y las instancias incorrectamente clasificadas (*Incorrectly Classified Instances*).

Ilustración 9 Resultados de la Clasificación



Fuente: Autores

El proceso anterior se realizó con cada uno de los algoritmos de las diferentes técnicas, en la *Tabla 5* se muestran los resultados que se obtuvieron.

Tabla 5 Evaluación de Algoritmos

ALGORITMO		INSTANCIAS CORRECTAMENTE CLASIFICADAS		INSTANCIAS INCORRECTAMENTE CLASIFICADAS	
GRUPO	NOMBRE	%	CANT.	%	CANT.
BAYES	BayesNet	57%	651	42,69%	485
BAYES	NaiveBayes	54,67%	621	45,33%	515
BAYES	NaiveBayesMultinomialText	23,59%	268	76,41%	868
BAYES	NaiveBayesUpdateable	54,67%	621	45,33%	515
TREES	DesicionStump	26,50%	301	73,50%	835
TREES	HoeffdingTree	57,48%	653	42,52%	483
TREES	J48	72,18%	820	27,82%	316
TREES	RandomTree	98,50%	1119	1,49%	17
TREES	REPTree	57,57%	654	42,43%	482

Fuente: Autores

Con base en los resultados obtenidos de la ejecución de cada uno de los algoritmos por técnica, se decidió continuar con el clasificador “*RandomTree*” el cual se encuentra en el grupo de algoritmos de la técnica de árboles de decisión “*TREES*”, ya que este es quien tiene el mayor porcentaje de instancias correctamente clasificadas, es decir, que el árbol de decisión que se generó tiene un nivel de confianza en la clasificación de 98,50%.

2.5. TIPOS DE CÁNCER CON MAYOR INCIDENCIA Y PRINCIPALES CAUSAS DIRECTAS DE MUERTE

Finalmente, con la etapa de pre-proceso realizado en WEKA, se creó la *Tabla 6*, donde se muestran los tipos de cáncer que tuvieron más incidencia dentro del periodo de tiempo analizado.

Tabla 6 Tipos de Cáncer con Mayor Incidencia

TIPO DE CANCER		INCIDENCIA	
CODIGO CIE10	NOMBRE	# DE CASOS	% DE INCIDENCIA
C349	TUMOR MALIGNO DE LOS BRONQUIOS O DEL PULMON	181	9,92%
C189	TUMOR MALIGNO DEL COLON	180	9,86%
C169	TUMOR MALIGNO DEL ESTOMAGO	163	8,93%
C509	TUMOR MALIGNO DE LA MAMA	100	5,48%
C61X	TUMOR MALIGNO DE LA PROSTATA	92	5,04%
C229	TUMOR MALIGNO DEL HIGADO	82	4,49%
C80X	TUMOR MALIGNO DE SITIOS NO ESPECIFICADOS	73	4,00%
D381	TUMOR DE COMPORTAMIENTO INCIERTO O DESCONOCIDO DE LA TRAQUEA, DE LOS BRONQUIOS Y DEL PULMON	68	3,73%
C159	TUMOR MALIGNO DEL ESOFAGO	58	3,18%
C900	MIELOMA MULTIPLE	53	2,90%
C762	TUMOR MALIGNO DEL ABDOMEN	51	2,79%
C449	TUMOR MALIGNO DE LA PIEL, SITIO NO ESPECIFICADO	35	1,92%

Fuente: Autores

De la misma manera, se creó la *Tabla 7* donde se muestran las principales causas de muerte ocasionadas a partir del tipo de cáncer que padecía la persona fallecida.

Tabla 7 Principales Causas de Muerte

TIPO DE CANCER		CAUSA DIRECTA DE MUERTE		
CÓDIGO CIE10	NOMBRE	CODIGO CIE10	NOMBRE	% DE INCIDENCIA POR CANCER
C349	TUMOR MALIGNO DE LOS BRONQUIOS O DEL PULMON	R570	CHOQUE CARDIOGENICO	18%
		J960	INSUFICIENCIA RESPIRATORIA AGUDA	12%
		I469	PARO CARDIACO, NO ESPECIFICADO	10%
		R090	ASFIXIA	10%
C189	TUMOR MALIGNO DEL COLON	I219	INFARTO AGUDO DEL MIOCARDIO, SIN OTRA ESPECIFICACION	37%
		R570	CHOQUE CARDIOGENICO	14%
		R99X	OTRAS CAUSAS MAL DEFINIDAS Y LAS NO ESPECIFICADAS DE MORTALIDAD	14%
		I469	PARO CARDIACO, NO ESPECIFICADO	7%
C169	TUMOR MALIGNO DEL ESTOMAGO	R571	CHOQUE HIPOVOLEMICO	19%
		I469	PARO CARDIACO, NO ESPECIFICADO	13%
		R688	OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	9%
		J969	INSUFICIENCIA RESPIRATORIA, NO ESPECIFICADA	7%
C509	TUMOR MALIGNO DE LA MAMA	J969	INSUFICIENCIA RESPIRATORIA, NO ESPECIFICADA	16%
		A419	SEPTICEMIA, NO ESPECIFICADA	15%
		R99X	OTRAS CAUSAS MAL DEFINIDAS Y LAS NO ESPECIFICADAS DE MORTALIDAD	12%
		I469	PARO CARDIACO, NO ESPECIFICADO	11%
C61X	TUMOR MALIGNO DE LA PROSTATA	I219	INFARTO AGUDO DEL MIOCARDIO, SIN OTRA ESPECIFICACION	24%
		I469	PARO CARDIACO, NO ESPECIFICADO	14%
		R99X	OTRAS CAUSAS MAL DEFINIDAS Y LAS NO ESPECIFICADAS DE MORTALIDAD	13%
		R570	CHOQUE CARDIOGENICO	11%
C229	TUMOR MALIGNO DEL HIGADO	I219	INFARTO AGUDO DEL MIOCARDIO, SIN OTRA ESPECIFICACION	59%
		R99X	OTRAS CAUSAS MAL DEFINIDAS Y LAS NO ESPECIFICADAS DE MORTALIDAD	20%

		J969	INSUFICIENCIA RESPIRATORIA, NO ESPECIFICADA	9%
		I469	PARO CARDIACO, NO ESPECIFICADO	4%
C80X	TUMOR MALIGNO DE SITIOS NO ESPECIFICADOS	I469	PARO CARDIACO, NO ESPECIFICADO	30%
		J969	INSUFICIENCIA RESPIRATORIA, NO ESPECIFICADA	18%
		R092	PARO RESPIRATORIO	16%
		R688	OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	12%
D381	TUMOR DE COMPORTAMIENTO INCIERTO O DESCONOCIDO DE LA TRAQUEA, DE LOS BRONQUIOS Y DEL PULMON	I219	INFARTO AGUDO DEL MIOCARDIO, SIN OTRA ESPECIFICACION	84%
		R092	PARO RESPIRATORIO	12%
		J969	INSUFICIENCIA RESPIRATORIA, NO ESPECIFICADA	3%
		J189	NEUMONIA, NO ESPECIFICADA	1%
C159	TUMOR MALIGNO DEL ESOFAGO	I219	INFARTO AGUDO DEL MIOCARDIO, SIN OTRA ESPECIFICACION	98%
		R99X	OTRAS CAUSAS MAL DEFINIDAS Y LAS NO ESPECIFICADAS DE MORTALIDAD	41%
		I469	PARO CARDIACO, NO ESPECIFICADO	26%
		R688	OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	16%
C900	MIELOMA MULTIPLE	R571	CHOQUE HIPOVOLEMICO	30%
		J969	INSUFICIENCIA RESPIRATORIA, NO ESPECIFICADA	21%
		N189	INSUFICIENCIA RENAL CRONICA, NO ESPECIFICADA	13%
		I469	PARO CARDIACO, NO ESPECIFICADO	11%
C762	TUMOR MALIGNO DEL ABDOMEN	I219	INFARTO AGUDO DEL MIOCARDIO, SIN OTRA ESPECIFICACION	78%
		R99X	OTRAS CAUSAS MAL DEFINIDAS Y LAS NO ESPECIFICADAS DE MORTALIDAD	16%
		I10X	HIPERTENSION ESENCIAL (PRIMARIA)	4%
		I469	PARO CARDIACO, NO ESPECIFICADO	2%
C449	TUMOR MALIGNO DE LA PIEL, SITIO NO ESPECIFICADO	I219	INFARTO AGUDO DEL MIOCARDIO, SIN OTRA ESPECIFICACION	49%
		R688	OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	49%
		E149	DIABETES MELLITUS, NO ESPECIFICADA SIN MENCION DE COMPLICACION	49%
		I469	PARO CARDIACO, NO ESPECIFICADO	49%

Fuente: Autores

2.6. FACTORES QUE INCIDEN EN EL RIESGO DE CONTRAER CÁNCER

A continuación, en la *Tabla 8* se muestran de manera detallada los factores que inciden en el riesgo de contraer cáncer, con la particularidad de que estos factores pueden ser evitados.

Tabla 8 Factores que inciden en el riesgo de contraer cáncer y que se pueden evitar

TIPO DE CÁNCER QUE PUEDE PADECER	ALCOHOL	OBESIDAD	RADIACIÓN	PRODUCTOS QUÍMICOS	TABAQUISMO
TUMOR MALIGNO DE LA MAMA (mujer)_ PARTE NO ESPECIFICADA	1	1	1	0	0
TUMOR MALIGNO DE LOS BRONQUIOS O DEL PULMON_ PARTE NO ESPECIFICADA	0	0	1	1	1
TUMOR MALIGNO DE LA PROSTATA	0	0	0	0	0
TUMOR MALIGNO DEL ESTOMAGO_ PARTE NO ESPECIFICADA	1	0	0	1	1
TUMOR MALIGNO DEL HIGADO_ NO ESPECIFICADO	1	1	0	1	0
TUMOR MALIGNO DEL COLON_ PARTE NO ESPECIFICADA	0	1	0	0	1
TUMOR MALIGNO DEL ESOFAGO	1	1	0	0	1

TUMOR MALIGNO DE LA PIEL_ SITIO NO ESPECIFICADO	0	0	0	1	0
---	---	---	---	---	---

Fuente: Autores con base en información obtenida en las páginas del Instituto Nacional del Cáncer (2015) y American Society of Clinical Oncology (ASCO), (2019)

Seguidamente, en la *Tabla 9*, al igual que en la anterior tabla, se muestra factores que inciden en el riesgo de contraer cáncer, con la particularidad de que estos factores no se pueden evitar a excepción de los anticonceptivos orales y la luz solar.

Tabla 9 Factores que inciden en el riesgo de contraer cáncer y que no se pueden evitar

TIPO DE CÁNCER	ANTECEDENTES PERSONALES	ANTECEDENTES FAMILIARES
TUMOR MALIGNO DE LA MAMA (mujer)_ PARTE NO ESPECIFICADA	ANTICONCEPTIVOS ORALES - MESTRUACION TEMPRANA - CANCER DE MAMA - CANCER DE OVARIO	CANCER DE MAMA - CANCER DE OVARIO
TUMOR MALIGNO DE LOS BRONQUIOS O DEL PULMON_ PARTE NO ESPECIFICADA	VIH	CÁNCER DE PULMÓN
TUMOR MALIGNO DE LA PROSTATA	N/A	CÁNCER DE PRÓSTATA
TUMOR MALIGNO DEL ESTOMAGO_ PARTE NO ESPECIFICADA	HELICOBACTER PYLORI - ANEMIA PERNICIOSA - ACLORIHIDRIA - CIRUGIA DE ESTOMAGO	CANCER DE MAMA - CANCER DE OVARIO CÁNCER DE ESTÓMAGO - CÁNCER GÁSTRICO
TUMOR MALIGNO DEL HIGADO_ NO ESPECIFICADO	HEPATITIS - HEPATITIS VIRAL - CIRROSIS - VIH	N/A
TUMOR MALIGNO DEL COLON_ PARTE NO ESPECIFICADA	ENFERMEDAD INFLAMATORIA INTESTINAL - CANCER COLORRECTAL - CANCER DE OVARIO - CANCER DE UTERO -	CÁNCER DE COLON
TUMOR MALIGNO DEL ESOFAGO	CANCER DE PULMON - CANCER DE BOCA - CANCER DE GARGANTA - VIRUS DEL PAPILOMA HUMANO – REFLUJO	N/A

TUMOR MALIGNO DE LA PIEL_ SITIO NO ESPECIFICADO	ENFERMEDADES EN LA PIEL - VIH - LUZ SOLAR - LUNARES ATÍPICOS	CÁNCER DE PIEL
---	--	----------------

Fuente: Autores con base en información obtenida en las páginas del Instituto Nacional del Cáncer (2015) y American Society of Clinical Oncology (ASCO), (2019)

3. ANALIZAR LOS RESULTADOS CON BASE EN LA IDENTIFICACIÓN DE PATRONES O COMPORTAMIENTOS DE DATOS DE INTERÉS.

En el diagnóstico ejecutado en el objetivo anterior, se obtuvo un árbol de decisión que se muestra en el *Anexo 1*, dicho árbol, arrojado por el clasificador “*RandomTree*”, cuenta con un conjunto de condiciones que se deben cumplir, para así llegar a un resultado final; este árbol inicia en la raíz, la cual indica una condición inicial, y finaliza al llegar a las hojas, que indican el resultado.

3.1. ESTRUCTURA DEL ÁRBOL DE DECISIONES

En este caso, la estructura del árbol de decisiones se encuentra de la siguiente manera:

3.1.1. Raíz

Es representada por la causa anticipada (C_ANT1), allí inicia el conjunto de condiciones. La causa anticipada, es la que predomina ante los demás atributos de acuerdo a los resultados arrojados por el software.

3.1.2. Nodos internos

Hacen referencia a los posibles caminos que pueden seguir, partiendo desde la raíz. Los nodos internos del árbol de decisiones obtenido son representados por la edad (EDAD), el sexo (SEXO) y la causa base (C_BASE).

3.1.3. Hojas

Son el resultado final del árbol de decisiones, se representa por las causas directas de muerte (C_DIRECTA). Debido a que el estudio se centra en el análisis de la mortalidad, en el software se tomó como etiqueta la causa directa de muerte de la persona.

3.2. MATRIZ DE CONFUSIÓN

Por otro lado, el algoritmo de árbol de decisión arrojó una matriz de confusión, ver *Anexo 2*. Esta matriz permite visualizar cómo está clasificando los datos el algoritmo; cuando las instancias correctamente clasificadas representan el total de los datos, quiere decir que la matriz estará compuesta por una diagonal que va

- # de instancias por causa real. Contiene la cantidad de instancias incorrectamente clasificadas respecto a la causa directa real.

Tabla 10 Instancias incorrectamente clasificadas

CAUSA DIRECTA ASIGNADA	# DE INSTANCIAS		CAUSA DIRECTA REAL	# DE INSTANCIAS POR CAUSA REAL
	CORRECTAMENTE CLASIFICADAS	INCORRECTAMENTE CLASIFICADAS		
J969	74	2	J960	1
			R99X	1
R570	70	6	I469	2
			R99X	2
			I219	1
			R69X	1
I469	125	1	J81X	1
I219	267	3	I679	1
			I674	1
			K922	1
J960	33	1	R571	1
R092	46	2	R99X	2
R571	61	1	R092	1
I678	7	1	I609	1

Fuente: Autores

3.3. APLICACIÓN PARA EL USO DE RESULTADOS

Tomando como referencia el proyecto realizado por *Calleja (2010)* y teniendo en cuenta la finalidad del presente estudio, se decidió realizar una aplicación web para dar uso a los resultados obtenidos. Esta aplicación permitió crear una predicción de la posible causa de muerte directa que tendría un paciente que padece alguno de los tipos de cáncer más frecuentes mostrados en la *Tabla 6*.

3.3.1. Atributos que contiene la aplicación

- *Sexo*: Masculino o femenino
- *Edad*: Edad del paciente a quien se realizará la predicción.
- *Causa base*: Tipo de cáncer que padece.
- *Causa anticipada*: Algún antecedente del paciente a quien se realizará la predicción.

El motivo por el cual se decidió crear la aplicación de predicción, es el de brindar una herramienta de apoyo en los hospitales que son quienes deben trabajar por mantener la salud de sus usuarios, no solo atendiendo sus emergencias, sino

también previniéndolas; a través de la predicción de la posible causa de muerte que tendría el paciente, se pueden tomar planes de acción para evitar dichas causas y así, buscar la prolongación de la vida de aquellos pacientes que padecen alguna de estas enfermedades que tanto inciden en la mortalidad.

Para programar la aplicación se tuvo como guía el proyecto antes mencionado, el cual, plantea como trabajos futuros realizar aplicaciones de predicción en aplicativos webs para facilitar el uso y el acceso. La programación se realizó en PHP, el cual “es un lenguaje de código abierto muy popular especialmente adecuado para el desarrollo web y que puede ser incrustado en HTML.” (PHP Group, 2019). A continuación, se muestra la ilustración 11, donde se muestra la interfaz del formulario de predicción.

Ilustración 11 Interfaz de formulario de predicción

CODIGO	CAUSA BASE	CODIGO	CAUSA ANTICIPADA
C159	TUMOR MALIGNO DEL ESOFAGO, PARTE NO ESPECIFICADA	A09X	DIARREA Y GASTROENTERITIS DE PRESUNTO ORIGEN INFECCIOSO
C169	TUMOR MALIGNO DEL ESTOMAGO, PARTE NO ESPECIFICADA	A153	TUBERCULOSIS DEL PULMON, CONFIRMADA POR MEDIOS NO ESPECIFICADOS
C189	TUMOR MALIGNO DEL COLON, PARTE NO ESPECIFICADA	A162	TUBERCULOSIS DE PULMON, SIN MENCION DE CONFIRMACION BACTERIOLOGICA O HISTOLOGICA
C229	TUMOR MALIGNO DEL HIGADO, NO ESPECIFICADO	A170	MENINGITIS TUBERCULOSA (G01*)
C349	TUMOR MALIGNO DE LOS BRONQUIOS O DEL PULMON, PARTE NO ESPECIFICADA	A412	SEPTICEMIA DEBIDA A ESTAFILOCOCO NO ESPECIFICADO
C449	TUMOR MALIGNO DE LA PIEL, SITIO NO ESPECIFICADO	A415	SEPTICEMIA DEBIDA A OTROS ORGANISMOS GRAMNEGATIVOS
C509	TUMOR MALIGNO DE LA MAMA, PARTE NO ESPECIFICADA	A418	OTRAS SEPTICEMIAS ESPECIFICADAS
C61X	TUMOR MALIGNO DE LA PROSTATA	A419	SEPTICEMIA NO ESPECIFICADA

Fuente: Autores

3.3.2. Funcionamiento del aplicativo web

- Cuenta con una lista seleccionable para indicar el sexo del paciente (masculino o femenino).
- Un campo de texto para ingresar en formato numérico la edad del paciente.
- Dos campos de texto que también funcionan como buscadores, para ingresar la causa base y la causa anticipada del paciente.
- Un botón que aparece en la parte inferior del formulario llamado “PREDECIR”, el cual, al dar clic sobre él, buscará el resultado que corresponda a los datos de entrada.

3.3.3. Árbol de decisión en la programación

En la *Ilustración 12*, se muestra una parte de la programación que relaciona el árbol de decisiones con el formulario y finalmente el botón “PREDECIR”.

Ilustración 12 Ejemplo, programación relacionada con el árbol de decisión

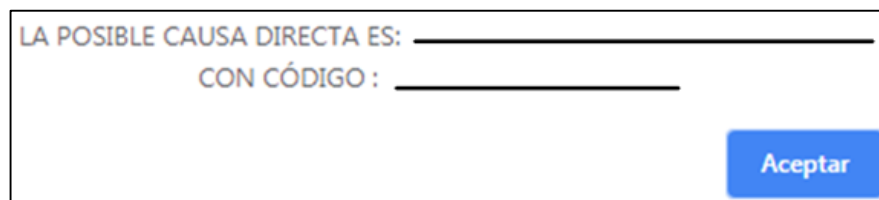
```
1 <?php
2
3 //Datos traídos del formulario
4
5 $SEXO_binarized = $_POST["sex"];
6 $EDAD = $_POST["edad"];
7 $C_BASE = $_POST["cba"];
8 $C_ANT1 = $_POST["ca1"];
9 $C_DIRECTA="";
10
11
12
13 if ($C_ANT1 == 'I269'){
14     if($C_BASE == 'C509' ){ $C_DIRECTA= 'J969';}
15     else if($C_BASE == 'C349'){
16         if ($EDAD < 76 ){ $C_DIRECTA= 'R570';}
17         else if ($EDAD >= 76){ $C_DIRECTA= 'J960';}
18     }
19     else if($C_BASE == 'C61X' ){ $C_DIRECTA= 'I219';}
20     else if($C_BASE == 'C169' ){ $C_DIRECTA= 'J969';}
21     else if($C_BASE == 'C80X'){
22         if ($SEXO_binarized < 0.5 ){ $C_DIRECTA= 'R092';}
23         else if ($SEXO_binarized >= 0.5 ){ $C_DIRECTA= 'J969';}
24     }
25     else if($C_BASE == 'C229' ){ $C_DIRECTA= 'R960';}
26     else if($C_BASE == 'C189'){
27         if ($EDAD < 91.5){
28             if ($EDAD < 73 ){ $C_DIRECTA= 'I219';}
29             else if ($EDAD >= 73){
30                 if ($SEXO_binarized < 0.5 ){ $C_DIRECTA= 'I219';}
```

Fuente: Autores

3.3.4. Resultado de la predicción

Al ingresar los datos en el formulario y dar clic sobre el botón “PREDECIR”, se activa la búsqueda del resultado en el árbol de decisiones, mediante el cumplimiento de los condicionales, hasta llegar a un resultado final, el cual es devuelto al usuario en forma de ventana emergente; ver *Ilustración 13*. Al dar clic en “Aceptar”, regresará automáticamente al formulario.

Ilustración 13 Resultado de la predicción



LA POSIBLE CAUSA DIRECTA ES: _____

CON CÓDIGO : _____

Aceptar

Fuente: Autores

Cabe aclarar que el formulario de predicción aplica solo para los tipos de cáncer seleccionados en el capítulo 2 como los más frecuentes dentro de la población objeto de estudio, ver *Tabla 6 (Tipos de Cáncer con Mayor Incidencia)*.

4. CONCLUSIONES

- De acuerdo a lo tratado en el objetivo 1, y teniendo en cuenta que la minería de datos es un proceso por el cual se puede obtener información que se encuentra implícita en los datos almacenados, con un sector tan importante como el de la salud se pueden llegar a obtener resultados que aporten ventajas a las entidades de dicho sector y de la misma manera apoyen a un mejor desempeño en pro de la salud de sus pacientes. Cada vez son más las organizaciones que se apoyan en la minería de datos buscando actuar de forma preventiva ante las situaciones de riesgo que se puedan presentar y así mejorar los diagnósticos, lo que a su vez puede verse reflejado en la cantidad de pacientes y la gravedad de sus casos.
- Así mismo, el sector salud es un campo donde a diario se obtienen grandes cantidades de datos las cuales se deben aprovechar para mejorar la calidad del servicio y crear una confianza en los usuarios, ya que muchas veces se abstienen de él por el hecho de que no les dan un diagnóstico confiable o suficiente para aportar una mejora en su estado de salud.
- En definitiva, la minería de datos no sería lo suficiente sin el apoyo de las diferentes herramientas que permiten su aplicación, Weka se ha convertido en una de las herramientas más utilizadas para el análisis de datos en el sector salud. La mayoría de los estudios tomados en cuenta en este proyecto han sido realizados con el apoyo de dicha herramienta debido a la facilidad de acceso y manejo con que se cuenta al utilizarla, también por la diversidad de métodos que contiene y la calidad de sus resultados.
- Con respecto al objetivo 2, finalmente se trabajaron con doce (12) tipos de cáncer, doscientos veinticinco (225) antecedentes de enfermedades que pudo haber tenido el paciente, ochenta y siete (87) posibles causas directas de muerte, personas de sexo femenino y masculino, y con edades que van hasta los noventa y nueve (99) años.
- El sector agrícola, siendo el principal sector económico en el Valle del Cauca, y de acuerdo a las diferentes actividades que se ejecutan en este sector, tales como aplicación de insumos y riegos a los cultivos, se llega a la conclusión de que dicho sector es un generador de factores que inciden en el riesgo de contraer cáncer. Las personas expuestas en este entorno (*Valle del Cauca*), tienen un riesgo de contraer diferentes tipos de cáncer, entre los cuales se encuentra: el cáncer de pulmón, cáncer de estómago y cáncer de piel.
- Las principales industrias en el Valle del Cauca como la azucarera, la producción de papel, químicos, maquinaria y equipos; al igual que el sector agrícola, en sus diferentes actividades se generan factores que hacen

vulnerables a las personas de contraer cáncer. En dichas industrias se utilizan una variedad de químicos o sustancias tóxicas que llegan a las personas de forma directa (*personas que aplican las sustancias*) o indirecta (*riego de cultivos con agua proveniente de los desechos industriales que se filtran en los ríos, lagos, entre otros*).

- Así pues, los diferentes factores generados en el Valle del Cauca, debido a las diferentes actividades que se llevan a cabo, pudieron ser las causas que condujeron a contraer los diferentes tipos de cáncer a las personas analizadas en la base de datos.
- El algoritmo RandomTree permitió asociar y clasificar los cinco (5) atributos con un 98,50% de precisión.
- Tras el análisis de los datos en el objetivo 3 y al obtener el árbol de decisiones, se pudo observar que el atributo predominante fue la causa anticipada (C_ANT1), ya que esta fue tomada por el algoritmo como la raíz del árbol de decisiones. La razón de esto puede ser la cantidad mayor de causas anticipadas con respecto a la cantidad de datos en los demás atributos o, desde el punto de vista médico, se debe a que las causas anticipadas son situaciones que se presentan a los pacientes dependiendo del tipo de enfermedad que padezcan, y factores como la edad y el sexo que influyen en la ocurrencia de dichas causas.
- Para el caso de la etiqueta o datos de salida del árbol de decisiones, se tomó la causa directa (C_DIRECTA) que es aquella causa que provoca la muerte de cada paciente dentro de la base de datos. Del análisis se obtuvo que la casusa directa con más incidencia fue el Infarto Agudo de Miocardio con código I219, el cual es generado por una disminución del flujo sanguíneo.
- La matriz de confusión obtenida del análisis de los datos, permitió la visualización de los errores en la clasificación. Estos errores disminuyen el nivel de efectividad del algoritmo; en este caso, se tendría un 1,50% de probabilidad de error en la clasificación de los datos de entrada, es decir, que la predicción realizada por la aplicación tendrá un 98,50% de confianza.
- A través de la aplicación web se logró dar un impacto positivo en el sector salud; con dicha aplicación, específicamente los médicos, podrán tomar planes de acción partiendo de la causa directa de muerte que posiblemente afectará el paciente, esto con base en la edad, sexo, tipo de cáncer que padece y alguna otra enfermedad relevante.

FUTUROS TRABAJOS

- Se espera ampliar la cantidad de datos realizando la solicitud a la totalidad de los municipios del Valle del Cauca, de tal manera que se puedan analizar un mayor número de tipos de cáncer.
- Realizar este tipo de proyectos teniendo en cuenta otro tipo de enfermedad que incida en una alta tasa de mortalidad.

5. BIBLIOGRAFÍA

- American Cancer Society. (02 de Abril de 2019). *American Cancer Society - Sobre el cáncer - Tipos de cáncer por orden alfabético*. Obtenido de <https://www.cancer.org/es/cancer/tipos-de-cancer.html>
- American Society of Clinical Oncology (ASCO). (08 de 04 de 2019). *Cancer.Net*. Obtenido de <https://www.cancer.net/cancer-types>
- Barra, J. M. (s.f.). *Microsystem*. Obtenido de <https://www.microsystem.cl/plataforma/rapidminer/>
- Calleja, A. J. (12 de 2010). *Riunet*. Obtenido de <https://riunet.upv.es/handle/10251/10097?show=full>
- Chekanov. (22 de Febrero de 2013). *jHepWork*. Obtenido de http://descargar.cnet.com/jHepWork/3000-2070_4-75833656.html
- Departamento del Valle del Cauca. (08 de 04 de 2019). *ENCOLOMBIA.COM - INICIO - GEOGRAFÍA - DEPARTAMENTOS COLOMBIANOS*. Obtenido de <https://encolombia.com/educacion-cultura/geografia/departamentos/valle-del-cauca/>
- DIGITAL, S. P. (s.f.). *Civic Software*. Obtenido de <https://www.softwarepublicocolombia.gov.co/en/public-or-civic-software/orange-data-mining>
- Dursun, D., Glenn, W., & Amit, K. (07 de 2005). *ScienceDirect*. Obtenido de <https://www.sciencedirect.com/science/article/pii/S0933365704001010>
- Garcia, L. (25 de Octubre de 2011). *UN POCO DE JAVA Y +*. Obtenido de <https://unpocodejava.com/2011/10/25/jhepwork-java-en-entornos-cientificos/>
- Guevara, V. (09 de 04 de 2019). *Scientific Databases - Home - Current Projects - Text Mining - Minería de Texto - WEKA*. Obtenido de <http://www.scientificdatabases.ca/current-projects/english-spanish-text-data-mining/mineria-de-texto-ingles-espanol/weka-waikato-environment-for-knowledge-analysis/>
- Instituto Nacional del Cáncer. (23 de Diciembre de 2015). *Instituto Nacional del Cáncer - Página principal - El cáncer - Causas y prevención*. Obtenido de <https://www.cancer.gov/espanol/cancer/causas-prevencion/riesgo>
- KNIME. (s.f.). Obtenido de <https://www.knime.com/knime-software>
- Liga Contra el Cáncer*. (s.f.). Obtenido de <http://www.ligacontraelcancer.com.co/cancer-de-cuello-uterino/>
- Ljubljana, U. d. (s.f.). *ORANGE*. Obtenido de <https://orange.biolab.si/features/visual-programming/>
- López, C. P. (2007). *Minería de Datos*. España: Thomson Ediciones Paraninfo S.A.
- Ministerio de Salud y Protección Social. (03 de Abril de 2019). *Ministerio de Salud y Protección Social - Salud - Salud Pública - Ministerio - Prevención enfermedades NO transmisibles - Cáncer*. Obtenido de <https://www.minsalud.gov.co/salud/publica/PENT/Paginas/Prevenciondel-cancer.aspx>
- Ministerio de Salud y Protección Social; Instituto Nacional de Cancerología. (Marzo de 2012). *Plan Nacional Para el Control del Cáncer en Colombia 2012-2020*.

- Obtenido de
<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/IA/INCA/plan-nacional-control-cancer-2012-2020.pdf>
- Organización Mundial de la Salud - OMS. (12 de Septiembre de 2018). *Sitio Web Mundial - Centro de Prensa: Cáncer*. Obtenido de <https://www.who.int/es/news-room/fact-sheets/detail/cancer>
- Organización Mundial de la Salud - OMS. (2 de Abril de 2019). *Sitio Web Mundial - Temas de Salud: Cáncer*. Obtenido de <http://www.who.int/topics/cancer/es/>
- Oztekin, A., Delen, D., & Kong, Z. (2009). Predicting the graft survival for heart–lung transplantation. *ELSEVIER*, 13.
- Perez, C., & Santín, D. (2008). *MINERÍA DE DATOS: TÉCNICAS Y HERRAMIENTAS*. MADRID, ESPAÑA.
- PHP Group. (10 de 04 de 2019). *php - Documentation*. Obtenido de <https://www.php.net/manual/es/intro-what-is.php>
- Salud, S. d. (2016). *Análisis de Situación de Salud 2016 del Municipio de Zarzal*.
- Salud, S. D. (2016). *Análisis de Situación de Salud con el Modelo de los Determinantes Sociales Valle del Cauca Secretaría Departamental De Salud del Valle del Cauca Grupo de Salud Pública Grupo ASIS*. Departamento del Valle del Cauca.
- Secretaría Departamental De Salud del Valle del Cauca. (12 de Diciembre de 2018). *GOBERNACIÓN DEL VALLE DEL CAUCA - ASIS Departamento del Valle del Cauca*. Obtenido de <https://www.valledelcauca.gov.co/documentos/10685/asis-departamento-del-valle-del-cauca/>
- Secretaría Departamental De Salud del Valle del Cauca. (06 de Agosto de 2018). *GOBERNACIÓN VALLE DEL CAUCA*. Obtenido de <https://www.valledelcauca.gov.co/documentos/10683/asis-municipios-2017/genPagDocs=3>
- Soni, J., Ansari, U., Sharma, D., & Soni, S. (03 de 2011). *International Journal of Computer Applications*. Obtenido de https://www.researchgate.net/publication/258493784_Predictive_Data_Mining_for_Medical_Diagnosis_An_Overview_of_Heart_Disease_Prediction
- Timarán, R., & Yépez, M. C. (2016). Caracterización de la supervivencia de mujeres con cáncer invasivo de cuello uterino usando minería de datos. *REVISTA DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN*, 13.
- Torey, S. (08 de 04 de 2019). *Nuestraesfera - tema: agua*. Obtenido de <http://nuestraesfera.cl/zoom/contaminacion-del-agua-por-actividades-industriales/>
- Universia Costa Rica. (4 de Septiembre de 2017). *Universia Costa Rica - Noticias - Educación - Tipos de investigación: Descriptiva, Exploratoria y Explicativa*. Obtenido de <http://noticias.universia.cr/educacion/noticia/2017/09/04/1155475/tipos-investigacion-descriptiva-exploratoria-explicativa.html>

Vega, C., Rosano, G., López, J., Cendejas, J., & Ferreira, H. (08 de 04 de 2019). *International Institute of Informatics and Systemics*. Obtenido de www.cds2012/cd2012sci/cisci_2012/paperspdf/ca732ov.pdf

Waikato, U. o. (s.f.). *Descargando e instalando Weka*. Obtenido de <https://www.cs.waikato.ac.nz/ml/weka/downloading.html>

Waikato, U. o. (s.f.). *Software de minería de datos en Java*. Obtenido de <https://www.cs.waikato.ac.nz/ml/weka/index.html>

6. ANEXOS

A continuación, se muestra el Anexo 1 que corresponde al árbol de decisiones, dividido en diez (10) partes debido a la cantidad de líneas que contiene. Cada parte cuenta con tres columnas que deben ser leídas de izquierda a derecha.

6.1. Anexo 1. ÁRBOL DE DECISIONES

PARTE 1

RandomTree	C_BASE = C229 : 3969 (0/0)	C_BASE = C349 : 3969 (0/0)
C_Ant1 = I269	C_BASE = C189 : R99X (2/0)	C_BASE = C61X : I469 (1/0)
C_BASE = C349	C_BASE = C159 : 3969 (0/0)	C_BASE = C169
EDAD < 76 : R570 (1/0)	C_BASE = C449 : 3969 (0/0)	EDAD < 71.5 : R688 (2/0)
EDAD >= 76 : 3960 (3/0)	C_BASE = D381 : 3969 (0/0)	EDAD >= 71.5 : 3969 (1/0)
C_BASE = C61X : I219 (2/0)	C_BASE = C762 : 3969 (0/0)	C_BASE = C80X : I469 (1/0)
C_BASE = C169 : 3969 (1/0)	C_BASE = C900 : A490 (1/0)	C_BASE = C229 : 3969 (0/0)
C_BASE = C80X	SEXO_binarized < 0.5 : I469 (1/0)	C_BASE = C189 : D649 (1/0)
SEXO_binarized < 0.5 : R092 (1/0)	C_Ant1 = D849	C_BASE = C159 : 3969 (0/0)
SEXO_binarized >= 0.5 : 3969 (3/0)	SEXO_binarized >= 0.5 : L089 (1/0)	C_BASE = C449 : 3969 (0/0)
C_BASE = C229 : R960 (1/0)	C_Ant1 = D381	C_BASE = D381 : 3969 (0/0)
C_BASE = C189	SEXO_binarized < 0.5 : R092 (1/0)	C_BASE = C762 : 3969 (0/0)
EDAD < 91.5	SEXO_binarized >= 0.5	C_BASE = C900 : 3969 (0/0)
EDAD < 73 : I219 (1/0)	EDAD < 72 : 3969 (1/0)	C_Ant1 = I10X
EDAD >= 73	EDAD >= 72 : J80X (1/0)	EDAD < 85.5
SEXO_binarized < 0.5 : I219 (1/0)	C_Ant1 = C189	C_BASE = C509 : 3969 (0/0)
SEXO_binarized >= 0.5 : R570 (1/0)	EDAD < 53.5	C_BASE = C349
EDAD >= 91.5 : R99X (1/0)	EDAD < 47.5 : K650 (1/0)	EDAD < 69.5
C_BASE = C159 : 3969 (0/0)	EDAD >= 47.5 : 3969 (1/0)	EDAD < 51.5 : I500 (1/0)
C_BASE = C449 : 3969 (1/0)	EDAD >= 53.5	EDAD >= 51.5 : L089 (1/0)
C_BASE = D381 : 3969 (0/0)	SEXO_binarized < 0.5	EDAD >= 69.5 : R570 (2/0)
C_BASE = C762 : 3969 (0/0)	EDAD < 79 : I219 (2/0)	C_BASE = C61X
C_BASE = C900 : 3969 (0/0)	EDAD >= 79 : R99X (1/0)	EDAD < 42.5 : N189 (1/0)
C_Ant1 = E43X	SEXO_binarized >= 0.5	EDAD >= 42.5 : I219 (4/0)
C_BASE = C509 : I269 (1/0)	EDAD < 64	C_BASE = C169 : I219 (2/0)
C_BASE = C349 : R570 (1/0)	EDAD < 59.5 : I469 (1/0)	C_BASE = C80X
C_BASE = C169 : R090 (1/0)	EDAD >= 59.5 : I219 (1/0)	EDAD < 78.5 : I678 (2/1)
C_BASE = C80X : 3969 (1/0)	EDAD >= 64	EDAD >= 78.5 : I678 (1/0)
C_BASE = C229 : 3969 (0/0)	EDAD < 71.5 : C790 (1/0)	C_BASE = C229
C_BASE = C189 : 3969 (0/0)	EDAD >= 71.5 : I469 (1/0)	EDAD < 75.5
C_BASE = C381 : 3969 (0/0)	C_Ant1 = I679	EDAD < 66.5 : I219 (4/0)
C_BASE = C900 : 3969 (0/0)	C_BASE = C509 : 3969 (0/0)	EDAD >= 66.5
C_Ant1 = E875 : I469 (1/0)	C_BASE = C349	SEXO_binarized < 0.5 : R99X (1/0)
C_Ant1 = I251	SEXO_binarized < 0.5 : I469 (1/0)	SEXO_binarized >= 0.5
C_BASE = C509 : I219 (2/0)	SEXO_binarized >= 0.5 : 3960 (1/0)	EDAD >= 73.5 : R99X (1/0)
C_BASE = C349	C_BASE = C169 : 3969 (0/0)	EDAD >= 75.5 : I219 (8/0)
EDAD < 78.5 : I319 (1/0)	C_BASE = C80X	C_BASE = C189
EDAD >= 78.5 : R570 (1/0)	EDAD < 77 : I469 (1/0)	EDAD < 82.5
C_BASE = C61X : I219 (5/0)	EDAD >= 77 : R688 (1/0)	SEXO_binarized < 0.5
C_BASE = C169 : I469 (1/0)	C_BASE = C229 : 3969 (0/0)	EDAD < 77 : I219 (3/0)
C_BASE = C80X : 3969 (0/0)	C_BASE = C189 : I219 (1/0)	EDAD >= 77
C_BASE = C229 : I219 (2/0)	C_BASE = C159 : 3969 (0/0)	EDAD < 80 : I219 (2/1)
C_BASE = C189 : I219 (3/0)	C_BASE = C449 : 3969 (0/0)	EDAD >= 80 : I219 (2/0)
C_BASE = C900 : 3969 (0/0)	C_BASE = D381 : 3969 (0/0)	SEXO_binarized >= 0.5
C_BASE = C229 : I219 (2/0)	C_BASE = C762 : 3969 (0/0)	EDAD < 55 : R99X (1/0)
C_BASE = C189 : I219 (3/0)	C_BASE = C900 : 3969 (0/0)	EDAD >= 55
C_BASE = C159 : 3969 (0/0)	C_Ant1 = I517 : I219 (1/0)	EDAD < 78.5 : I219 (4/0)
	C_Ant1 = R571	EDAD >= 78.5 : R570 (1/0)
	C_BASE = C509 : 3969 (0/0)	SEXO_binarized < 0.5
	C_BASE = C349 : 3969 (0/0)	

PARTE 2

EDAD < 83.5 : I469 (1/0)	EDAD >= 45.5	C_BASE = C169 : 3969 (0/0)
EDAD >= 83.5 : I219 (1/0)	EDAD < 70.5	C_BASE = C80X : I678 (1/0)
SEXO_binarized >= 0.5	EDAD < 58	C_BASE = C229 : 3969 (0/0)
EDAD < 84.5 : I219 (2/1)	EDAD < 51 : I469 (1/0)	C_BASE = C189 : 3969 (0/0)
EDAD >= 84.5 : R99X (1/0)	EDAD >= 51 : I219 (1/0)	C_BASE = C169
C_BASE = C159	EDAD >= 58	EDAD < 76 : I219 (1/0)
SEXO_binarized < 0.5 : I469 (2/0)	EDAD < 65.5 : C798 (1/0)	EDAD >= 76 : R69X (1/0)
SEXO_binarized >= 0.5 : R99X (1/0)	EDAD >= 65.5 : I469 (1/0)	C_BASE = C159
C_BASE = C449	EDAD >= 70.5 : 3969 (1/0)	EDAD < 44.5
SEXO_binarized < 0.5 : I219 (3/0)	C_Ant1 = K566	EDAD < 63 : R601 (1/0)
SEXO_binarized >= 0.5	EDAD < 50.5 : A419 (1/0)	EDAD >= 63 : N19X (1/0)
EDAD < 66 : I219 (2/0)	EDAD >= 50.5 : R571 (2/0)	C_BASE = C449 : 3969 (0/0)
EDAD >= 66	C_Ant1 = J960	C_BASE = D381 : 3969 (0/0)
EDAD < 70.5 : E149 (1/0)	C_BASE = C509 : 3969 (0/0)	C_BASE = C762 : I219 (1/0)
EDAD >= 70.5 : I219 (1/0)	C_BASE = C349 : R092 (3/0)	C_BASE = C900 : 3969 (0/0)
C_BASE = D381 : I219 (11/0)	C_BASE = C61X : R960 (1/0)	SEXO_binarized >= 0.5
C_BASE = C762 : I219 (12/0)	C_BASE = C169 : R688 (1/0)	C_BASE = C509 : 3969 (0/0)
C_BASE = C900	C_BASE = C80X : I469 (1/0)	C_BASE = C349 : 3969 (0/0)
SEXO_binarized < 0.5	C_BASE = C229 : 3969 (0/0)	C_BASE = C61X : R570 (1/0)
EDAD < 67.5 : N189 (1/0)	C_BASE = C189 : 3969 (0/0)	C_BASE = C169 : 3969 (0/0)
EDAD >= 67.5 : I469 (1/0)	C_BASE = C159 : 3969 (0/0)	C_BASE = C80X : 3969 (0/0)
SEXO_binarized >= 0.5 : N189 (2/0)	C_BASE = C449 : R688 (1/0)	C_BASE = C229 : 3969 (0/0)
EDAD >= 85.5	C_BASE = D381 : R092 (1/0)	C_BASE = C189 : 3969 (0/0)
C_BASE = C509 : 3969 (0/0)	C_BASE = C762 : I219 (1/0)	C_BASE = C159
C_BASE = C349	C_BASE = C900 : 3969 (0/0)	EDAD < 76.5 : R99X (1/0)
EDAD < 87.5 : J159 (1/0)	C_Ant1 = C780	EDAD >= 76.5
EDAD >= 87.5 : I509 (1/0)	C_BASE = C509	EDAD < 84 : I469 (1/0)
C_BASE = C61X : I219 (1/0)	EDAD < 51	EDAD >= 84 : N19X (1/0)
C_BASE = C169 : J189 (1/0)	EDAD < 36 : R090 (1/0)	C_BASE = C449 : R570 (1/0)
C_BASE = C80X : I679 (1/0)	EDAD >= 36 : G931 (1/0)	C_BASE = D381 : 3969 (0/0)
C_BASE = C229 : I219 (5/0)	EDAD >= 51	C_BASE = C762 : 3969 (0/0)
C_BASE = C189	EDAD < 67.5 : C782 (1/0)	C_BASE = C900 : 3969 (0/0)
EDAD < 87.5 : I219 (3/0)	EDAD >= 67.5	C_Ant1 = A499
EDAD >= 87.5	EDAD < 76.5 : R688 (1/0)	EDAD < 50.5 : R570 (1/0)
SEXO_binarized < 0.5	EDAD >= 76.5 : R579 (1/0)	EDAD >= 50.5 : R571 (1/0)
EDAD < 88.5 : I219 (2/1)	C_BASE = C349 : 3969 (0/0)	EDAD < 50 : E86X (1/0)
EDAD >= 88.5 : I219 (2/0)	C_BASE = C61X : 3969 (0/0)	EDAD >= 50
SEXO_binarized >= 0.5 : I219 (1/0)	C_BASE = C169 : 3969 (0/0)	C_BASE = C509 : 3969 (0/0)
C_BASE = C159 : 3969 (0/0)	C_BASE = C80X : 3969 (1/0)	C_BASE = C349 : 3969 (0/0)
C_BASE = C449	C_BASE = C229 : 3969 (0/0)	C_BASE = C61X : 3969 (0/0)
EDAD < 93 : E149 (1/0)	C_BASE = C189 : I469 (2/0)	C_BASE = C169 : 3969 (0/0)
EDAD >= 93 : I219 (1/0)	C_BASE = C159 : 3969 (0/0)	C_BASE = C80X : I469 (1/0)
C_BASE = D381 : I219 (7/0)	C_BASE = C449 : 3969 (0/0)	C_BASE = C229 : 3969 (0/0)
C_BASE = C762 : I219 (4/0)	C_BASE = D381 : 3969 (0/0)	C_BASE = C189 : I219 (1/0)
C_BASE = C900 : I258 (1/0)	C_BASE = C762 : 3969 (0/0)	C_BASE = C159 : 3969 (0/0)
C_Ant1 = K631 : K619 (1/0)	C_BASE = C900 : 3969 (0/0)	C_BASE = C449 : 3969 (0/0)
C_Ant1 = N61X : D65X (1/0)	C_Ant1 = N189	C_BASE = D381 : 3969 (0/0)
C_Ant1 = C509	SEXO_binarized < 0.5	C_BASE = C762 : 3969 (0/0)
EDAD < 45.5	C_BASE = C509 : 3969 (0/0)	C_BASE = C900 : N189 (1/0)
EDAD < 41 : R688 (1/0)	C_BASE = C349 : 3969 (0/0)	C_Ant1 = R570
EDAD >= 41 : C539 (1/0)	C_BASE = C61X : N189 (1/0)	

PARTE 3

C_BASE = C509 : I469 (1/0) C_BASE = C349 SEXO_binarized < 0.5 : I469 (1/0) SEXO_binarized >= 0.5 : I501 (1/0) C_BASE = C61X : J969 (0/0) C_BASE = C169 EDAD < 73 : I469 (2/0) EDAD >= 73 : J81X (1/0) C_BASE = C80X : I469 (1/0) C_BASE = C229 EDAD < 82 : J969 (1/0) EDAD >= 82 : I219 (1/0) C_BASE = C189 : I469 (1/0) C_BASE = C159 : R688 (1/0) C_BASE = C449 EDAD < 74 : I219 (1/0) EDAD >= 74 : R688 (1/0) C_BASE = D381 : J969 (0/0) C_BASE = C762 : J969 (0/0) C_BASE = C900 : J969 (0/0) C_ANTI = I219 C_BASE = C509 : I469 (1/0) C_BASE = C349 EDAD < 65 EDAD < 62.5 : R570 (5/0) EDAD >= 62.5 SEXO_binarized < 0.5 : R570 (2/0) SEXO_binarized >= 0.5 : R090 (1/0) EDAD >= 65 EDAD < 71 : I509 (2/0) EDAD >= 71 SEXO_binarized < 0.5 EDAD < 81 EDAD < 79.5 : R570 (2/1) EDAD >= 79.5 : R570 (2/1) EDAD >= 81 EDAD < 82.5 : J960 (1/0) EDAD >= 82.5 EDAD < 85 : I509 (1/0) EDAD >= 85 EDAD < 88 : I469 (1/0) EDAD >= 88 : I509 (1/0) SEXO_binarized >= 0.5 EDAD < 76 : R570 (1/0) EDAD >= 76 EDAD < 79 : I469 (1/0) EDAD >= 79 EDAD < 83 : R570 (1/0) EDAD >= 83 : R570 (2/1) C_BASE = C61X EDAD < 82.5	EDAD < 66.5 EDAD < 54.5 : R570 (2/1) EDAD >= 54.5 : R570 (2/0) EDAD >= 66.5 EDAD < 74 : I709 (1/0) EDAD >= 74 EDAD < 80 : R570 (1/0) EDAD >= 80 EDAD < 81.5 : R960 (1/0) EDAD >= 81.5 : R570 (1/0) EDAD >= 82.5 EDAD < 85 : I469 (2/0) EDAD >= 85 : I251 (1/0) C_BASE = C169 EDAD < 81 EDAD < 72 EDAD < 62.5 : I461 (1/0) EDAD >= 62.5 : I469 (2/0) EDAD >= 72 EDAD < 78 : R092 (1/0) EDAD >= 78 : I10X (1/0) EDAD >= 81 EDAD < 94 EDAD < 89.5 : I469 (3/0) EDAD >= 89.5 EDAD < 90.5 : E872 (1/0) EDAD >= 90.5 : I469 (3/0) C_BASE = C80X : I469 (3/0) C_BASE = C229 SEXO_binarized < 0.5 EDAD < 65.5 : I469 (1/0) EDAD >= 65.5 : R960 (1/0) SEXO_binarized >= 0.5 : R99X (3/0) C_BASE = C189 SEXO_binarized < 0.5 EDAD < 72 : R570 (7/0) EDAD >= 72 EDAD < 81.5 EDAD < 77.5 : R99X (1/0) EDAD >= 77.5 : I469 (1/0) EDAD >= 81.5 EDAD < 88.5 : R570 (2/0) EDAD >= 88.5 : R570 (2/1) SEXO_binarized >= 0.5 EDAD < 69 EDAD < 48.5 : I219 (1/0) EDAD >= 48.5 : I469 (2/0) EDAD >= 69 EDAD < 86.5 EDAD < 79.5 : R570 (1/0)	EDAD >= 86.5 EDAD < 88.5 : R99X (1/0) EDAD >= 88.5 : R570 (1/0) C_BASE = C159 : J969 (0/0) C_BASE = C449 : J969 (0/0) C_BASE = D381 : J969 (0/0) C_BASE = C762 : R99X (1/0) C_BASE = C900 : J969 (0/0) C_ANTI = C900 : I469 (1/0) C_ANTI = R048 SEXO_binarized < 0.5 : R090 (1/0) SEXO_binarized >= 0.5 : J969 (1/0) C_ANTI = R64X : I219 (1/0) C_ANTI = A162 EDAD < 67 : R092 (1/0) EDAD >= 67 : I500 (1/0) C_ANTI = E46X : I469 (1/0) C_ANTI = C782 C_BASE = C509 EDAD < 54 : J969 (1/0) EDAD >= 54 : I219 (1/0) C_BASE = C349 EDAD < 72 : I469 (1/0) EDAD >= 72 : R092 (1/0) C_BASE = C61X : J969 (0/0) C_BASE = C169 : J969 (0/0) C_BASE = C80X : R092 (1/0) C_BASE = C229 : J969 (0/0) C_BASE = C189 : J969 (0/0) C_BASE = C159 : J969 (0/0) C_BASE = C449 : J969 (0/0) C_BASE = D381 : J969 (0/0) C_BASE = C762 : J969 (0/0) C_BASE = C900 : J969 (0/0) C_ANTI = C798 C_BASE = C509 : J969 (0/0) C_BASE = C349 : J969 (0/0) C_BASE = C61X : R570 (1/0) C_BASE = C169 : J969 (0/0) C_BASE = C80X : R688 (1/0) C_BASE = C229 : J969 (0/0) C_BASE = C189 : J969 (0/0) C_BASE = C159 : J969 (0/0) C_BASE = C449 : I469 (1/0) C_BASE = D381 : J969 (0/0) C_BASE = C762 : J969 (0/0) C_BASE = C900 : J969 (0/0) C_ANTI = R090 C_BASE = C509 : I469 (1/0) C_BASE = C349 EDAD < 68.5 : J960 (1/0)
---	---	---

PARTE 4

EDAD >= 68.5 : J159 (1/0) C_BASE = C61X : J969 (0/0) C_BASE = C169 : R688 (1/0) C_BASE = C80X : J969 (0/0) C_BASE = C229 : J969 (0/0) C_BASE = C189 : R99X (1/0) C_BASE = C159 : R99X (1/0) C_BASE = C449 : J969 (0/0) C_BASE = D381 : R092 (1/0) C_BASE = C762 : J969 (0/0) C_BASE = C900 : J969 (0/0) C_ANTI = D648 : E86X (1/0) C_ANTI = I871 : J969 (1/0) C_ANTI = N390 C_BASE = C509 EDAD < 77 EDAD < 53.5 : R579 (1/0) EDAD >= 53.5 EDAD < 63 : R571 (2/0) EDAD >= 63 EDAD < 71 : A419 (1/0) EDAD >= 71 : R571 (1/0) EDAD >= 77 EDAD < 82.5 : J960 (1/0) EDAD >= 82.5 SEXO_binarized < 0.5 EDAD < 88.5 EDAD < 84.5 : A419 (1/0) EDAD >= 84.5 : E46X (1/0) EDAD >= 88.5 : A419 (2/0) SEXO_binarized >= 0.5 : N390 (1/0) C_BASE = C349 : A419 (3/0) C_BASE = C61X EDAD < 80 : J960 (1/0) EDAD >= 80 EDAD < 92.5 : N40X (1/0) EDAD >= 92.5 : R99X (1/0) C_BASE = C169 EDAD < 76.5 EDAD < 50.5 : R571 (1/0) EDAD >= 50.5 EDAD < 69.5 : N390 (1/0) EDAD >= 69.5 : R092 (1/0) EDAD >= 76.5 SEXO_binarized < 0.5 EDAD < 91.5 : J180 (1/0) EDAD >= 91.5 : R571 (1/0) SEXO_binarized >= 0.5 EDAD < 83.5 : R571 (1/0) EDAD >= 83.5 : E872 (1/0) C_BASE = C80X : J969 (1/0)	EDAD >= 68.5 : J159 (1/0) C_BASE = C61X : J969 (0/0) C_BASE = C169 : R688 (1/0) C_BASE = C80X : J969 (0/0) C_BASE = C229 : J969 (0/0) C_BASE = C189 : R99X (1/0) C_BASE = C159 : R99X (1/0) C_BASE = C449 : J969 (0/0) C_BASE = D381 : R092 (1/0) C_BASE = C762 : J969 (0/0) C_BASE = C900 : J969 (0/0) C_ANTI = D648 : E86X (1/0) C_ANTI = I871 : J969 (1/0) C_ANTI = N390 C_BASE = C509 EDAD < 77 EDAD < 53.5 : R579 (1/0) EDAD >= 53.5 EDAD < 63 : R571 (2/0) EDAD >= 63 EDAD < 71 : A419 (1/0) EDAD >= 71 : R571 (1/0) EDAD >= 77 EDAD < 82.5 : J960 (1/0) EDAD >= 82.5 SEXO_binarized < 0.5 EDAD < 88.5 EDAD < 84.5 : A419 (1/0) EDAD >= 84.5 : E46X (1/0) EDAD >= 88.5 : A419 (2/0) SEXO_binarized >= 0.5 : N390 (1/0) C_BASE = C349 : A419 (3/0) C_BASE = C61X EDAD < 80 : J960 (1/0) EDAD >= 80 EDAD < 92.5 : N40X (1/0) EDAD >= 92.5 : R99X (1/0) C_BASE = C169 EDAD < 76.5 EDAD < 50.5 : R571 (1/0) EDAD >= 50.5 EDAD < 69.5 : N390 (1/0) EDAD >= 69.5 : R092 (1/0) EDAD >= 76.5 SEXO_binarized < 0.5 EDAD < 91.5 : J180 (1/0) EDAD >= 91.5 : R571 (1/0) SEXO_binarized >= 0.5 EDAD < 83.5 : R571 (1/0) EDAD >= 83.5 : E872 (1/0) C_BASE = C80X : J969 (1/0)	C_BASE = C229 : R99X (1/0) C_BASE = C189 SEXO_binarized < 0.5 : I219 (1/0) SEXO_binarized >= 0.5 : R99X (1/0) C_BASE = C159 : R99X (2/0) C_BASE = C449 : J969 (0/0) C_BASE = D381 : J969 (0/0) C_BASE = C762 : J969 (0/0) C_BASE = C900 : R571 (1/0) C_ANTI = K650 C_BASE = C509 : R571 (1/0) C_BASE = C349 : R570 (1/0) C_BASE = C61X : K650 (1/0) C_BASE = C169 : R571 (1/0) C_BASE = C80X : J969 (0/0) C_BASE = C229 : J969 (0/0) C_BASE = C189 : I219 (1/0) C_BASE = C159 : R688 (1/0) C_BASE = C449 : J969 (0/0) C_BASE = D381 : J969 (0/0) C_BASE = C762 : J969 (0/0) C_BASE = C900 SEXO_binarized < 0.5 : J969 (1/0) SEXO_binarized >= 0.5 : R571 (2/0) C_ANTI = K768 : D689 (1/0) C_ANTI = I509 C_BASE = C509 EDAD < 43 : J960 (1/0) EDAD >= 43 : I469 (1/0) C_BASE = C349 SEXO_binarized < 0.5 : R570 (1/0) SEXO_binarized >= 0.5 EDAD < 89 : R090 (1/0) EDAD >= 89 : R092 (1/0) C_BASE = C61X EDAD < 81.5 : I710 (1/0) EDAD >= 81.5 : I219 (2/0) C_BASE = C169 : J969 (0/0) C_BASE = C80X EDAD < 85 : R688 (1/0) EDAD >= 85 : I679 (1/0) C_BASE = C229 : I219 (2/0) C_BASE = C189 EDAD < 71 EDAD < 63.5 : I219 (2/0) EDAD >= 63.5 : R570 (2/1) EDAD >= 71 EDAD < 79.5 : E86X (1/0) EDAD >= 79.5 : I219 (1/0) C_BASE = C159 : I469 (1/0) C_BASE = C449 : J969 (0/0)
--	--	--

PARTE 5

C.BASE = D381 : I219 (3/0) C.BASE = C762 : I219 (2/0) C.BASE = C900 : J969 (0/0) C.ANT1 = R11X : R571 (1/0) C.ANT1 = B690 : I469 (1/0) C.ANT1 = I469 C.BASE = C509 : I469 (1/0) C.BASE = C349 EDAD < 55.5 : J80X (1/0) EDAD >= 55.5 EDAD < 83.5 : I469 (1/0) EDAD >= 83.5 : R570 (1/0) C.BASE = C61X EDAD < 63 : R960 (1/0) EDAD >= 63 : I469 (2/0) C.BASE = C169 EDAD < 75.5 : R571 (2/0) EDAD >= 75.5 SEXO_binarized < 0.5 : R99X (1/0) SEXO_binarized >= 0.5 EDAD < 83 : E878 (1/0) EDAD >= 83 : I259 (1/0) C.BASE = C80X EDAD < 70.5 EDAD < 57.5 : I469 (1/0) EDAD >= 57.5 : R092 (2/0) EDAD >= 70.5 EDAD < 79 : R688 (1/0) EDAD >= 79 : I469 (3/0) C.BASE = C229 : R99X (2/0) C.BASE = C189 : R99X (1/0) C.BASE = C159 EDAD < 74.5 : R99X (1/0) EDAD >= 74.5 EDAD < 83 : R688 (1/0) EDAD >= 83 : R99X (1/0) C.BASE = C449 : R688 (1/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : R99X (1/0) C.BASE = C900 : J969 (0/0) C.ANT1 = I749 SEXO_binarized < 0.5 EDAD < 78.5 : I679 (1/0) EDAD >= 78.5 : I219 (1/0) SEXO_binarized >= 0.5 : R570 (1/0) C.ANT1 = C20X : R579 (1/0) C.ANT1 = L984 : R092 (1/0) C.ANT1 = A419 C.BASE = C509 : J969 (0/0) C.BASE = C349 : J960 (1/0) C.BASE = C61X : J969 (0/0)	C.BASE = C169 : E872 (1/0) C.BASE = C80X : J969 (0/0) C.BASE = C229 : J969 (0/0) C.BASE = C189 : J969 (0/0) C.BASE = C159 : J969 (0/0) C.BASE = C449 : J969 (0/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : E872 (3/0) C.ANT1 = D489 : R570 (1/0) C.ANT1 = D649 SEXO_binarized < 0.5 : R571 (1/0) SEXO_binarized >= 0.5 : I678 (1/0) C.ANT1 = L089 C.BASE = C509 : A419 (1/0) C.BASE = C349 EDAD < 61.5 : L089 (1/0) EDAD >= 61.5 : J960 (1/0) C.BASE = C61X : J969 (0/0) C.BASE = C169 : R571 (2/1) C.BASE = C80X : J969 (0/0) C.BASE = C229 : J969 (0/0) C.BASE = C189 : J969 (0/0) C.BASE = C159 : J969 (0/0) C.BASE = C449 : J969 (0/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : R571 (1/0) C.ANT1 = J81X C.BASE = C509 : R99X (1/0) C.BASE = C349 SEXO_binarized < 0.5 : J80X (1/0) SEXO_binarized >= 0.5 : J960 (2/0) C.BASE = C61X : I469 (1/0) C.BASE = C169 SEXO_binarized < 0.5 EDAD < 70 : R99X (1/0) EDAD >= 70 : I219 (1/0) SEXO_binarized >= 0.5 : R092 (2/0) C.BASE = C80X : J969 (1/0) C.BASE = C229 : J969 (1/0) C.BASE = C189 : R99X (1/0) C.BASE = C159 : R99X (1/0) C.BASE = C449 : J969 (0/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : J969 (1/0) C.ANT1 = E149 C.BASE = C509 : J969 (0/0) C.BASE = C349 EDAD < 75 : I269 (1/0)	EDAD >= 75 : R092 (1/0) C.BASE = C61X : N189 (1/0) C.BASE = C169 : I10X (2/0) C.BASE = C80X : J969 (0/0) C.BASE = C229 : R99X (2/0) C.BASE = C189 SEXO_binarized < 0.5 : D649 (1/0) SEXO_binarized >= 0.5 EDAD < 71 : I219 (1/0) EDAD >= 71 EDAD < 81.5 : R99X (1/0) EDAD >= 81.5 : I219 (1/0) C.BASE = C159 : R99X (1/0) C.BASE = C449 EDAD < 71 : I219 (1/0) EDAD >= 71 SEXO_binarized < 0.5 : R570 (1/0) SEXO_binarized >= 0.5 : E141 (1/0) C.BASE = D381 : I219 (3/0) C.BASE = C762 EDAD < 65.5 : R99X (1/0) EDAD >= 65.5 : I219 (2/0) C.BASE = C900 : N189 (1/0) C.ANT1 = C795 : I219 (1/0) C.ANT1 = I319 EDAD < 70 : R090 (1/0) EDAD >= 70 : G931 (1/0) C.ANT1 = G938 C.BASE = C509 : J969 (0/0) C.BASE = C349 : R090 (1/0) C.BASE = C61X : N189 (1/0) C.BASE = C169 : J969 (0/0) C.BASE = C80X : J969 (0/0) C.BASE = C229 : J969 (0/0) C.BASE = C189 : J969 (0/0) C.BASE = C159 : J969 (0/0) C.BASE = C449 : J969 (0/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : J969 (0/0) C.ANT1 = E80X C.BASE = C509 : J969 (0/0) C.BASE = C349 : J969 (0/0) C.BASE = C61X : J969 (0/0) C.BASE = C169 : I469 (1/0) C.BASE = C80X : R688 (1/0) C.BASE = C229 : J969 (0/0) C.BASE = C189 : J969 (0/0) C.BASE = C159 : J969 (0/0) EDAD < 78.5 : J80X (1/0) EDAD >= 78.5 : R99X (1/0) C.BASE = C159 : J969 (0/0)
--	--	--

PARTE 6

C.BASE = C449 : J969 (0/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : J969 (0/0) C.ANT1 = B690 C.BASE = C509 : J969 (1/0) C.BASE = C349 : J969 (0/0) C.BASE = C61X : J969 (0/0) C.BASE = C169 SEXO_binarized < 0.5 : J81X (1/0) SEXO_binarized >= 0.5 : R13X (1/0) C.BASE = C80X EDAD < 83.5 : J984 (1/0) EDAD >= 83.5 : R092 (1/0) C.BASE = C229 : J969 (0/0) C.BASE = C189 : N390 (1/0) C.BASE = C159 : I469 (1/0) C.BASE = C449 : J969 (0/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : J969 (0/0) C.ANT1 = L039 C.BASE = C509 : R99X (1/0) C.BASE = C349 : L089 (1/0) C.BASE = C61X : J969 (0/0) C.BASE = C169 : J969 (0/0) C.BASE = C80X : J969 (0/0) C.BASE = C229 : J969 (0/0) C.BASE = C189 : J969 (0/0) C.BASE = C159 : R99X (1/0) C.BASE = C449 : J969 (0/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : J969 (0/0) C.ANT1 = G979 : R99X (1/0) C.ANT1 = R899 EDAD < 75 : R99X (2/0) EDAD >= 75 : R69X (1/0) C.ANT1 = R092 C.BASE = C509 : A419 (1/0) C.BASE = C349 EDAD < 87 EDAD < 79.5 : R090 (1/0) EDAD >= 79.5 EDAD < 84 : J960 (1/0) EDAD >= 84 : L089 (1/0) EDAD >= 87 SEXO_binarized < 0.5 : R092 (1/0) SEXO_binarized >= 0.5 : R090 (2/0) C.BASE = C61X : J969 (0/0) C.BASE = C169 : J969 (0/0)	C.BASE = C80X : J969 (0/0) C.BASE = C229 : J969 (0/0) C.BASE = C189 : J969 (0/0) C.BASE = C159 : J969 (0/0) C.BASE = C449 : J969 (0/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : I469 (1/0) C.ANT1 = R042 : R99X (1/0) C.ANT1 = R579 C.BASE = C509 : J969 (0/0) C.BASE = C349 EDAD < 29.5 : J960 (1/0) EDAD >= 29.5 : I469 (1/0) C.BASE = C61X : J969 (0/0) C.BASE = C169 : E872 (1/0) C.BASE = C80X : R688 (1/0) C.BASE = C229 : I518 (1/0) C.BASE = C189 : R99X (1/0) C.BASE = C159 : R688 (1/0) C.BASE = C449 : R688 (1/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : J969 (0/0) C.ANT1 = R351 : I469 (1/0) C.ANT1 = R18X : R99X (1/0) C.ANT1 = R99X : I469 (3/0) C.ANT1 = R578 C.BASE = C509 : J969 (0/0) C.BASE = C349 : J969 (0/0) C.BASE = C61X : J969 (0/0) C.BASE = C169 : J969 (0/0) C.BASE = C80X : J969 (0/0) C.BASE = C229 : J969 (1/0) C.BASE = C189 : J969 (0/0) C.BASE = C159 : J969 (0/0) C.BASE = C449 : R688 (1/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : J969 (0/0) C.ANT1 = D619 EDAD < 66 : R571 (1/0) EDAD >= 66 : E872 (1/0) C.ANT1 = I420 C.BASE = C509 : J969 (0/0) C.BASE = C349 : R001 (1/0) C.BASE = C61X : J969 (0/0) C.BASE = C169 : I219 (2/0) C.BASE = C80X : J969 (0/0) C.BASE = C229 : J969 (0/0) C.BASE = C189 : J969 (0/0)	C.BASE = C159 : J969 (0/0) C.BASE = C449 : I219 (1/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 : J969 (0/0) C.ANT1 = N179 C.BASE = C509 : R578 (1/0) C.BASE = C349 : J969 (0/0) C.BASE = C61X : R99X (1/0) C.BASE = C169 : J969 (0/0) C.BASE = C80X : J969 (0/0) C.BASE = C229 EDAD < 45 : I219 (1/0) EDAD >= 45 EDAD < 86 : R99X (1/0) EDAD >= 86 : I219 (1/0) C.BASE = C189 : J969 (0/0) C.BASE = C159 : J969 (0/0) C.BASE = C449 : J969 (0/0) C.BASE = D381 : J969 (0/0) C.BASE = C762 : J969 (0/0) C.BASE = C900 EDAD < 37 : R571 (1/0) EDAD >= 37 : E870 (1/0) C.ANT1 = I500 C.BASE = C509 : J969 (0/0) C.BASE = C349 : R570 (1/0) C.BASE = C61X : J969 (0/0) C.BASE = C169 SEXO_binarized < 0.5 : I219 (1/0) SEXO_binarized >= 0.5 EDAD < 73.5 : I219 (1/0) EDAD >= 73.5 EDAD < 80 : I10X (1/0) EDAD >= 80 : R092 (1/0) C.BASE = C80X : I469 (1/0) C.BASE = C229 : I219 (1/0) C.BASE = C189 EDAD < 81.5 EDAD < 63.5 : I219 (1/0) EDAD >= 63.5 : R570 (1/0) EDAD >= 81.5 : I219 (2/0) C.BASE = C159 : I469 (1/0) C.BASE = C449 : J969 (0/0) C.BASE = C381 : J969 (0/0) C.BASE = C762 : I219 (3/0) C.BASE = C900 : I469 (1/0) C.ANT1 = A419 C.BASE = C509 EDAD < 68.5 : R99X (1/0) EDAD >= 68.5
---	---	--

PARTE 7

EDAD < 82 : A419 (1/0) EDAD >= 82 : N390 (1/0) C_BASE = C349 EDAD < 24.5 : R090 (1/0) EDAD >= 24.5 SEXO_binarized < 0.5 : I269 (1/0) SEXO_binarized >= 0.5 : R090 (1/0) C_BASE = C61X : 1960 (2/0) C_BASE = C169 SEXO_binarized < 0.5 EDAD < 82 : R571 (1/0) EDAD >= 82 : R688 (1/0) SEXO_binarized >= 0.5 : I189 (2/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1219 (1/0) C_BASE = C189 EDAD < 47 : D689 (1/0) EDAD >= 47 EDAD < 78.5 : R100 (1/0) EDAD >= 78.5 : I469 (1/0) C_BASE = C159 EDAD < 81.5 : I469 (1/0) EDAD >= 81.5 : R99X (3/0) C_BASE = C449 : 1969 (0/0) C_BASE = C381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 SEXO_binarized < 0.5 : I469 (1/0) SEXO_binarized >= 0.5 : 1969 (1/0) C_ANT1 = F328 : 1960 (1/0) C_ANT1 = S271 : 1969 (3/0) C_ANT1 = I119 C_BASE = C509 : 1219 (1/0) C_BASE = C349 : 1500 (1/0) C_BASE = C61X : 1219 (1/0) C_BASE = C169 : 1461 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C219 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = T659 C_BASE = C509 : G931 (1/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : 1969 (1/0) C_BASE = C229 : 1969 (0/0)	C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = C381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (1/0) C_ANT1 = I209 : 1219 (1/0) C_ANT1 = G919 : 1619 (1/0) C_ANT1 = I499 C_BASE = C509 : R99X (1/0) C_BASE = C349 : R090 (3/0) C_BASE = C61X : 1219 (1/0) C_BASE = C169 : 1259 (1/0) C_BASE = C80X : 1469 (1/0) C_BASE = C229 : R99X (1/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1469 (1/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : R99X (1/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = T175 C_BASE = C509 : 1969 (2/0) C_BASE = C349 : 1960 (1/0) C_BASE = C61X : 1960 (1/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : 1969 (2/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : R80X (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = I255 C_BASE = C509 : 1219 (1/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1219 (3/0) C_BASE = C189 : R99X (1/0) C_BASE = C159 SEXO_binarized < 0.5 : R99X (1/0) SEXO_binarized >= 0.5 : R688 (1/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1219 (1/0) C_BASE = C762 : 1219 (1/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = I259 EDAD < 84 C_BASE = C509 : 1969 (0/0)	C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1219 (1/0) C_BASE = C80X : 1469 (1/0) C_BASE = C229 : 1219 (1/0) C_BASE = C189 : 1219 (3/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1219 (1/0) C_BASE = D381 : 1219 (4/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) EDAD >= 84 : R99X (1/0) C_ANT1 = I249 : 1219 (1/0) C_ANT1 = I519 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1989 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1219 (1/0) C_BASE = C189 : 1219 (2/0) C_BASE = C449 : 1219 (2/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = I429 SEXO_binarized < 0.5 : R99X (1/0) SEXO_binarized >= 0.5 : I219 (4/0) C_ANT1 = J449 EDAD < 86 C_BASE = C509 EDAD < 75.5 : 1969 (1/0) EDAD >= 75.5 EDAD < 79.5 : R99X (1/0) EDAD >= 79.5 : 1969 (1/0) C_BASE = C349 EDAD < 76.5 EDAD < 65 : J159 (2/0) EDAD >= 65 EDAD < 69.5 : I469 (1/0) EDAD >= 69.5 : 1961 (1/0) EDAD >= 76.5 EDAD < 80.5 : R092 (1/0) EDAD >= 80.5 EDAD < 83 : R090 (1/0) EDAD >= 83 SEXO_binarized < 0.5 : R090 (1/0) SEXO_binarized >= 0.5 : J159 (1/0) C_BASE = C61X : 1960 (1/0) C_BASE = C169
--	---	---

PARTE 8

SEXO_binarized < 0.5 EDAD < 54 : J180 (1/0) EDAD >= 54 : J189 (1/0) SEXO_binarized >= 0.5 EDAD < 82.5 : J189 (2/0) EDAD >= 82.5 : R688 (1/0) C_BASE = C80X : 1969 (1/0) C_BASE = C219 : 1969 (0/0) EDAD < 61 : 1969 (1/0) EDAD >= 61 EDAD < 82 : I219 (1/0) EDAD >= 82 : J981 (1/0) C_BASE = C189 SEXO_binarized < 0.5 : I219 (1/0) SEXO_binarized >= 0.5 : J80X (1/0) C_BASE = C159 : R99X (1/0) C_BASE = C449 : 1219 (1/0) C_BASE = D381 : 1219 (1/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1469 (1/0) EDAD >= 86 C_BASE = C509 : 1969 (1/0) C_BASE = C349 EDAD < 89.5 : R060 (1/0) EDAD >= 89.5 SEXO_binarized < 0.5 EDAD < 94.5 : R092 (1/0) EDAD >= 94.5 : R570 (1/0) SEXO_binarized >= 0.5 : R092 (1/0) C_BASE = C61X : R99X (1/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : 1969 (0/0) C_BASE = C219 : 1969 (0/0) C_BASE = C189 : 1219 (2/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1219 (1/0) C_BASE = C900 : N189 (1/0) C_ANT1 = I619 C_BASE = C509 : 1969 (1/0) C_BASE = C349 : R570 (2/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : R571 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C219 : 1969 (0/0) C_BASE = C189 : 1619 (1/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0)	C_BASE = C900 : 1969 (0/0) C_ANT1 = A418 C_BASE = C509 : R99X (2/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : R99X (2/0) C_BASE = C169 : J81X (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C219 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : R99X (5/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = Q249 : I509 (1/0) C_ANT1 = J984 C_BASE = C509 : 1969 (1/0) C_BASE = C349 EDAD < 77.5 : A419 (1/0) EDAD >= 77.5 : 1960 (1/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : R092 (1/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : E149 (1/0) C_BASE = D381 : 1219 (1/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (1/0) C_ANT1 = I694 C_BASE = C509 EDAD < 78.5 : R578 (1/0) EDAD >= 78.5 : E46X (1/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : K746 (1/0) C_BASE = C169 SEXO_binarized < 0.5 : J984 (1/0) SEXO_binarized >= 0.5 : N390 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : 1219 (1/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1219 (1/0) C_BASE = D381 : 1219 (1/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = R54X EDAD < 89 : I679 (1/0) EDAD >= 89 : I219 (1/0) C_ANT1 = J80X	EDAD < 71 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : J81X (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C219 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : R092 (1/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (1/0) EDAD >= 71 SEXO_binarized < 0.5 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : R571 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : R092 (1/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) SEXO_binarized >= 0.5 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : R090 (2/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : R99X (1/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = J869 SEXO_binarized < 0.5 : R571 (1/0) SEXO_binarized >= 0.5 : I269 (1/0) C_ANT1 = E440 : I219 (1/0) C_ANT1 = J441 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : 1960 (1/0) C_BASE = C61X : 1469 (1/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1969 (0/0)
---	---	--

PARTE 9

C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (1/0) C_ANT1 = L89X C_BASE = C509 : R99X (1/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (0/0) EDAD < 69 : R688 (1/0) EDAD >= 69 : R571 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : A499 (1/0) C_ANT1 = I447 : I219 (1/0) C_ANT1 = I678 C_BASE = C509 : R578 (1/0) C_BASE = C349 : I269 (1/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = K830 EDAD < 84.5 : K650 (1/0) EDAD >= 84.5 : R571 (1/0) C_ANT1 = S273 SEXO_binarized < 0.5 : 1969 (1/0) SEXO_binarized >= 0.5 : D62X (1/0) C_ANT1 = S269 : 1969 (1/0) C_ANT1 = I139 SEXO_binarized < 0.5 : R99X (1/0) SEXO_binarized >= 0.5 : I509 (1/0) C_ANT1 = A09X C_BASE = C509 : 1969 (0/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : K922 (1/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : 1969 (0/0)	C_BASE = C229 : 1969 (0/0) C_BASE = C189 : K650 (1/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : E878 (1/0) C_ANT1 = R568 : I619 (1/0) C_ANT1 = I64X EDAD < 67.5 : I469 (2/0) EDAD >= 67.5 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : R092 (3/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = T792 : D62X (2/0) C_ANT1 = J440 : J189 (1/0) C_ANT1 = I501 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : I678 (1/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = K769 : K650 (1/0) C_ANT1 = E872 EDAD < 71.5 : E872 (2/0) EDAD >= 71.5 : I469 (2/0) C_ANT1 = T71X : 1969 (3/0) C_ANT1 = R31X : D689 (1/0) C_ANT1 = T751 : 1969 (1/0) C_ANT1 = K565 : K650 (1/0) C_ANT1 = R100 : K650 (1/0) C_ANT1 = I609 : I619 (2/0) C_ANT1 = S361 : D62X (1/0) C_ANT1 = E876 : I219 (1/0) C_ANT1 = R262 : 1969 (1/0)	C_ANT1 = E142 : I219 (1/0) C_ANT1 = I48X EDAD < 72.5 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : R570 (1/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : I678 (1/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : R99X (1/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : I219 (1/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) EDAD >= 72.5 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : I509 (1/0) C_BASE = C61X : K650 (1/0) C_BASE = C169 : 1969 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : I219 (1/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : I219 (1/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = E785 : I219 (2/0) C_ANT1 = R986 : 1969 (1/0) C_ANT1 = T909 : E46X (1/0) C_ANT1 = T600 : 1969 (1/0) C_ANT1 = G309 EDAD < 86 : J159 (1/0) EDAD >= 86 : I690 (1/0) C_ANT1 = K85X EDAD < 24 : R688 (1/0) EDAD >= 24 : K650 (1/0) C_ANT1 = G301 EDAD < 86.5 : R688 (1/0) EDAD >= 86.5 : J189 (1/0) C_ANT1 = F03X SEXO_binarized < 0.5 : I678 (1/0) SEXO_binarized >= 0.5 : E46X (1/0) C_ANT1 = A412 : A419 (1/0) C_ANT1 = M623 : A419 (1/0) C_ANT1 = T07X : 1969 (1/0) C_ANT1 = I340 SEXO_binarized < 0.5 : I219 (1/0) SEXO_binarized >= 0.5 : I251 (1/0) C_ANT1 = I633 : I219 (1/0)
--	--	--

PARTE 10

C_ANT1 = K550 : K650 (1/0) C_ANT1 = K746 : J81X (1/0) C_ANT1 = I718 : I219 (1/0) C_ANT1 = I714 : I219 (1/0) C_ANT1 = I428 : I219 (2/0) C_ANT1 = I698 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : J159 (1/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : 1969 (0/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : I219 (1/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : R688 (1/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : I219 (1/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = I339 : I219 (1/0) C_ANT1 = J180 EDAD < 75.5 : A419 (2/0) EDAD >= 75.5 : 1969 (1/0) SEXO_binarized >= 0.5 : 1969 (1/0) C_ANT1 = K269 : I10X (1/0) C_ANT1 = I454 : I259 (1/0) C_ANT1 = K659 EDAD < 70 : K650 (1/0) EDAD >= 70 : R571 (1/0) C_ANT1 = Z958 : I251 (1/0) C_ANT1 = I739 : I219 (1/0) C_ANT1 = R104 : K650 (1/0) C_ANT1 = Z988 : I710 (1/0) C_ANT1 = E789 : I219 (1/0) C_ANT1 = N19X EDAD < 78 : R571 (1/0) EDAD >= 78 SEXO_binarized < 0.5 : I469 (1/0) SEXO_binarized >= 0.5 : N390 (1/0) C_ANT1 = T318 : A419 (1/0) C_ANT1 = Q210 : I330 (1/0) C_ANT1 = S729 : I469 (1/0) C_ANT1 = J90X : 1969 (1/0) C_ANT1 = T179 : 1969 (1/0) C_ANT1 = I442 EDAD < 68 : R570 (1/0) EDAD >= 68 : I509 (1/0) C_ANT1 = E880 : I469 (1/0) C_ANT1 = G418 : I469 (1/0) C_ANT1 = I639 : I469 (1/0) C_ANT1 = S298 : 1969 (1/0)	C_ANT1 = J384 : J80X (1/0) C_ANT1 = I700 : I219 (1/0) C_ANT1 = G409 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : R568 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : 1969 (0/0) C_BASE = C159 : R99X (1/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = K833 : R571 (1/0) C_ANT1 = A170 : R571 (1/0) C_ANT1 = I693 : R688 (1/0) C_ANT1 = D62X : R571 (1/0) C_ANT1 = E109 SEXO_binarized < 0.5 : R688 (1/0) SEXO_binarized >= 0.5 : N189 (1/0) C_ANT1 = T794 : R092 (1/0) C_ANT1 = K767 : N19X (1/0) C_ANT1 = I490 EDAD < 46.5 : R99X (1/0) EDAD >= 46.5 : R570 (2/0) C_ANT1 = B379 : R571 (1/0) C_ANT1 = N12X : R571 (2/0) C_ANT1 = J22X : R571 (1/0) C_ANT1 = I959 : R570 (1/0) C_ANT1 = K259 : K922 (2/0) C_ANT1 = T314 : R688 (1/0) C_ANT1 = T145 : R58X (1/0) C_ANT1 = T093 : N390 (1/0) C_ANT1 = I634 : R578 (1/0) C_ANT1 = I921 : R571 (1/0) C_ANT1 = I210 : R570 (1/0) C_ANT1 = Q793 : Q249 (1/0) C_ANT1 = I711 : R570 (1/0) C_ANT1 = S299 : R092 (1/0) C_ANT1 = G936 : R090 (1/0) C_ANT1 = T814 : R688 (1/0) C_ANT1 = I713 : R571 (1/0) C_ANT1 = I350 : R570 (1/0) C_ANT1 = R851 : R571 (1/0) C_ANT1 = G912 : R568 (1/0) C_ANT1 = T011 : R092 (1/0) C_ANT1 = I258 : R570 (1/0) C_ANT1 = J041 : R571 (1/0) C_ANT1 = S099 : R092 (1/0) C_ANT1 = E141 : R571 (1/0) C_ANT1 = E145 : L089 (1/0) C_ANT1 = K729 : R688 (1/0) C_ANT1 = G934 : R99X (1/0) C_ANT1 = I951 : R99X (1/0) C_ANT1 = I719 C_BASE = C509 : 1969 (0/0) C_BASE = C349 : 1969 (0/0) C_BASE = C61X : 1969 (0/0) C_BASE = C169 : R571 (1/0) C_BASE = C80X : 1969 (0/0) C_BASE = C229 : 1969 (0/0) C_BASE = C189 : R101 (1/0) C_BASE = C159 : 1969 (0/0) C_BASE = C449 : 1969 (0/0) C_BASE = D381 : 1969 (0/0) C_BASE = C762 : 1969 (0/0) C_BASE = C900 : 1969 (0/0) C_ANT1 = K703 : R99X (1/0) C_ANT1 = N170 : R99X (1/0) C_ANT1 = K559 : R99X (1/0) C_ANT1 = I472 : R99X (1/0) C_ANT1 = M329 : R99X (1/0) C_ANT1 = R060 : R99X (1/0) C_ANT1 = I710 : R571 (1/0) C_ANT1 = F009 : R99X (1/0) C_ANT1 = T887 : R99X (1/0) C_ANT1 = A153 : R99X (1/0)
---	--

Fuente: Autores con base en el resultado arrojado por el software WEKA

A continuación, se muestra el Anexo 2 que corresponde a la matriz de confusión, dividido en cuatro (4) partes debido al número de columnas y filas que contiene.

6.2. Anexo 2. MATRIZ DE CONFUSIÓN

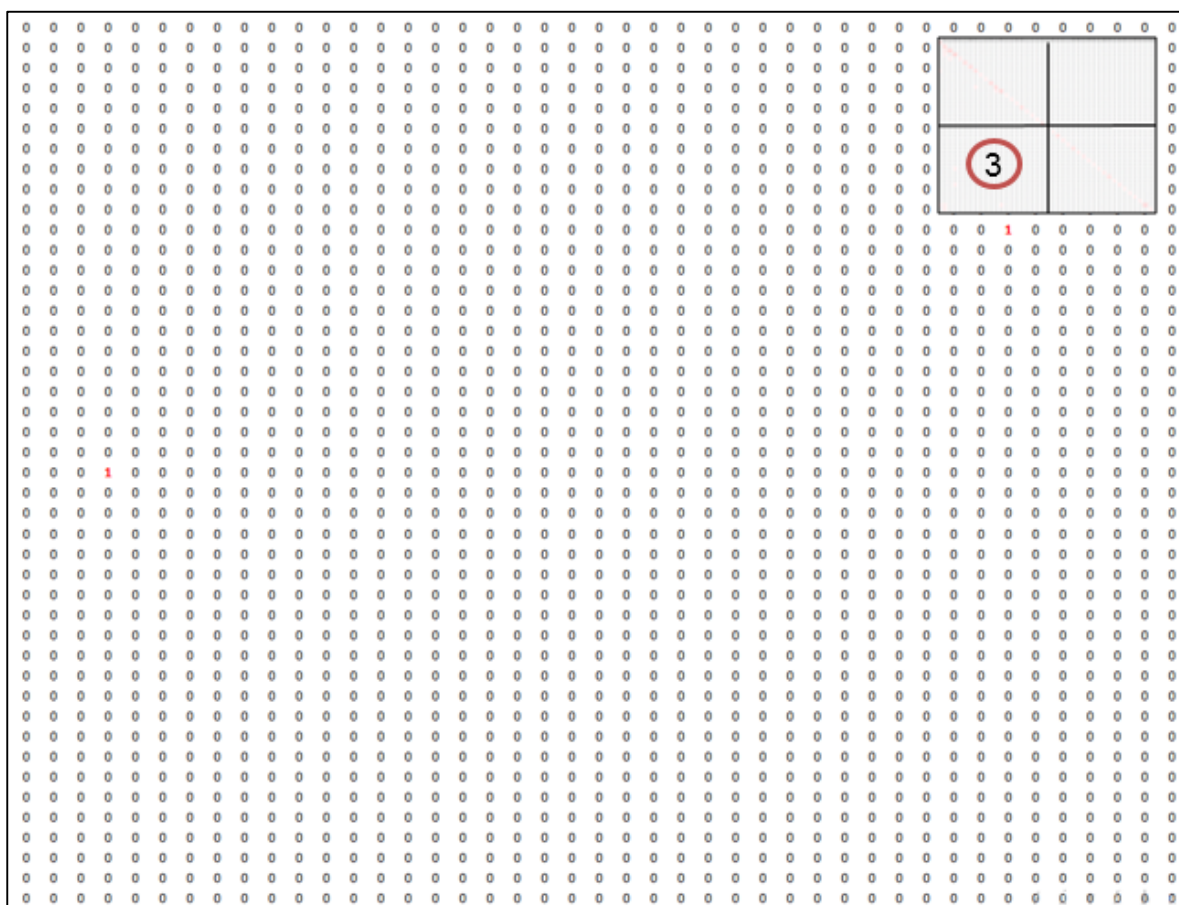
PARTE 1

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq						
a	74	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b	0	70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
c	0	0	2	125	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
d	0	0	1	0	267	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
e	0	0	0	1	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	0	0	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
g	0	0	0	0	0	0	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
h	1	0	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
i	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
j	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
k	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	0	0	0	0	0	0	0	1	0	0	0	61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	0	0	0	0	1	46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
o	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
r	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
t	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
w	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
z	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
aa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ab	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ac	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ad	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
af	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ag	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ah	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ai	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
aj	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ak	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
al	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
am	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
an	0	0	0	0	0																																												

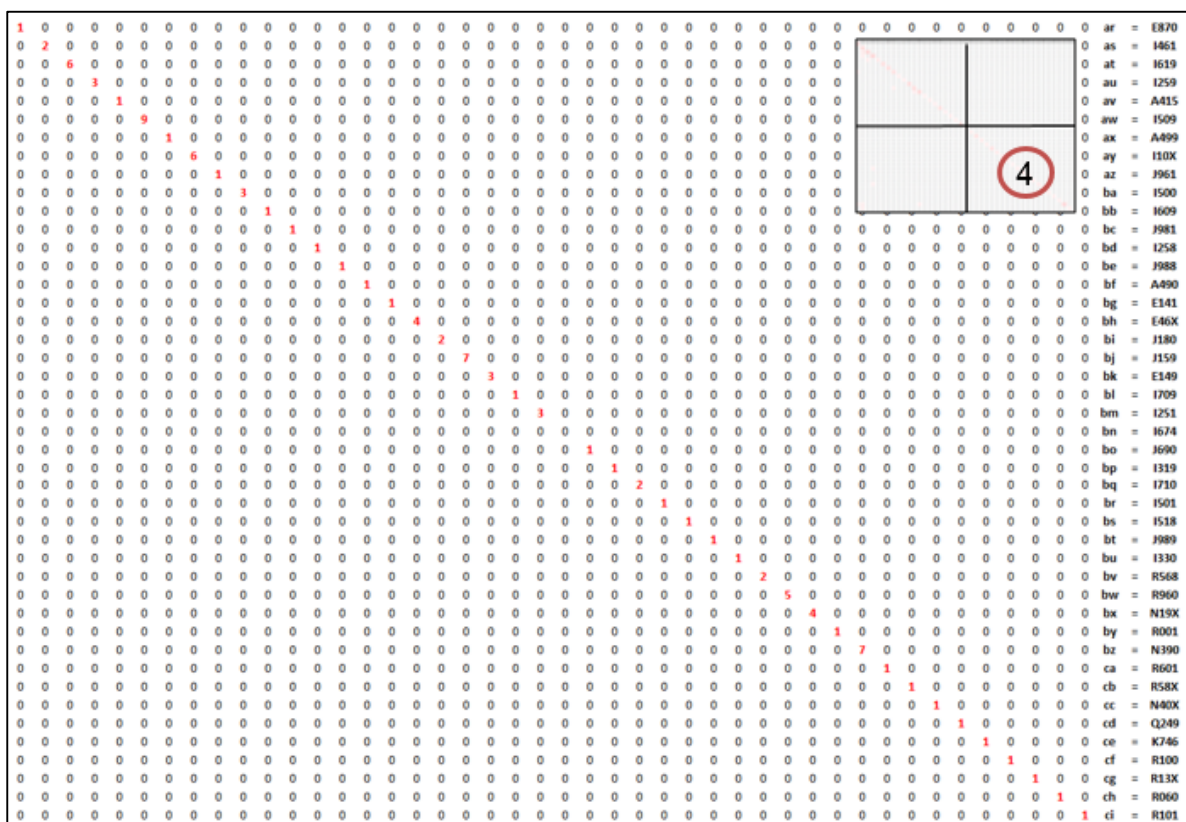
PARTE 2

	ar	as	at	au	av	aw	ax	ay	az	ba	bb	bc	bd	be	bf	bg	bh	bi	bj	bk	bl	bm	bn	bo	bp	bq	br	bs	bt	bu	bv	bw	bx	by	bz	ca	cb	cc	cd	ce	cf	cg	ch	ci	<
	a	=	J969																																										
	b	=	R570																																										
	c	=	A469																																										
	d	=	I219																																										
	e	=	X81X																																										
	f	=	R688																																										
	g	=	A419																																										
	h	=	J960																																										
	i	=	R57B																																										
	j	=	D62K																																										
	k	=	J189																																										
	l	=	R57I																																										
	m	=	K09Z																																										
	n	=	I269																																										
	o	=	K059																																										
	p	=	D65X																																										
	q	=	N189																																										
	r	=	C78Z																																										
	s	=	Z730																																										
	t	=	D619																																										
	u	=	J80V																																										
	v	=	R579																																										
	w	=	C798																																										
	x	=	R090																																										
	y	=	E86X																																										
	z	=	C780																																										
	aa	=	K050																																										
	ab	=	L089																																										
	ac	=	E878																																										
	ad	=	D649																																										
	ae	=	J984																																										
	af	=	H679																																										
	ag	=	G931																																										
	ah	=	E87Z																																										
	ai	=	K729																																										
	aj	=	K92Z																																										
	ak	=	H678																																										
	al	=	C790																																										
	am	=	O471																																										
	an	=	C539																																										
	ao	=	R99X																																										
	ap	=	R66X																																										
	aq	=	D64A																																										

PARTE 3



PARTE 4



Fuente: Autores con base en el resultado arrojado por el software WEKA