

# MINERÍA DE DATOS DE LA SALUD: SISTEMA DE VOTACIÓN DE TÉCNICAS ANALÍTICAS PARA IDENTIFICAR LOS FACTORES QUE INFLUYEN EN LA REALIZACIÓN DE CIRUGÍAS ESTÉTICAS

HEALTH DATA MINING SYSTEM: VOTING SYSTEM OF ANALYTICAL TECHNIQUES TO IDENTIFY THE FACTORS THAT INFLUENCE IN THE PERFORMANCE OF AESTHETIC SURGERIES

Ana Isabel Oviedo Carrascal <sup>1</sup>, Sandra Milena Sánchez Sánchez <sup>2</sup>

<sup>1</sup> Doctora en Ingeniería Electrónica – Énfasis en Descubrimiento de Conocimiento, Profesora titular en la Facultad de Ingeniería en Tecnologías de la Información y la Comunicación, Grupo de Investigación GIDATI. Correo electrónico: [ana.oviedo@upb.edu.co](mailto:ana.oviedo@upb.edu.co). Universidad Pontificia Bolivariana, circular 1 No. 70-01, Medellín, Colombia.

<sup>2</sup> Ingeniera Informática, Especialista en Inteligencia de Negocios, Estudiante de Maestría en Tecnologías de la Información y la Comunicación en Universidad Pontificia Bolivariana, Docente de Cátedra en la Facultad de Ingeniería Informática del Politécnico Colombiano Jaime Isaza Cadavid Restrepo. Correo electrónico: [sandra\\_sanchez89101@elpoli.edu.co](mailto:sandra_sanchez89101@elpoli.edu.co). Carrera 48 No. 7- 151, Medellín, Colombia.

## RESUMEN

A nivel mundial, las cirugías estéticas se han convertido en procedimientos de constante realización que requieren, como paso indispensable, la asistencia a una cita de valoración, pero muchas de estas citas no culminan con un procedimiento quirúrgico. Para identificar los factores que influyen en la realización de cirugías estéticas, en este artículo se propone un modelo analítico basado en un sistema de votación de varias técnicas como análisis de correlaciones, análisis de componentes principales, árboles de decisión, regresión logística y reglas de asociación. El modelo analítico es validado con los datos provenientes de un consultorio quirúrgico de la ciudad de Medellín. Los resultados indican que los pacientes que se realizan una cirugía estética son aquellos que: 1) manifiestan desear una segunda cirugía; 2) la primera cirugía deseada es facial, aumento mamario, bichectomía, abdominoplastia, aumento de implantes y glúteos; 3) el índice de masa corporal es normal; 4) no tienen hábitos de cigarrillo; y 5) no tiene antecedentes familiares de diabetes.

**Palabras clave:** Minería de datos de la Salud; Analítica de Datos; Ciencia de Datos.

Recibido: 19 de abril de 2017.

Received: April 19, 2017.

Aceptado: 30 de Junio de 2017.

Accepted: June 30, 2017.

## ABSTRACT

*At the global level, cosmetic surgeries have become constant procedures that require, as an indispensable step, attendance at an appointment, but many of these appointments do not culminate with a surgical procedure. In order to identify the factors that influence the performance of aesthetic surgeries, this article proposes an analytical model based on a voting system of several techniques such as correlation analysis, principal component analysis, decision trees, logistic regression and association rules. The analytical model is validated with data from a surgical office in the city of Medellín. The results indicate that patients who undergo a cosmetic surgery are those who: 1) claim to desire a second surgery; 2) the first surgery desired is facial, breast augmentation, bichectomy, abdominoplasty, implant and gluteal augmentation; 3) the body mass index is normal; 4) they do not have cigarette habits; And 5) have no family history of diabetes.*

**Keywords:** Data Mining Health; Data Analytics; Data Science.

*Cómo citar este artículo:* A. I. Oviedo-Carrascal, S. M. Sánchez-Sánchez, "Minería de datos de la salud: Sistema de votación de técnicas analíticas para identificar los factores que influyen en la realización de cirugías estéticas," *Revista Politécnica*, vol. 13, no. 25, pp. 43-52, 2017.

## 1. INTRODUCCIÓN

Las cirugías estéticas han venido en aumento en los últimos años [1], así lo reportan constantemente revistas como “The Wall Street Journal”, “The New York Times”, “The Boston Globe”, “Canadian Business”, entre otras [2]. En Colombia también se ha generado un aumento de citas de valoración con especialistas en cirugía plástica por parte de personas que desean mejorar su aspecto físico [3]. En los consultorios, el número de pacientes se incrementa no sólo por la experiencia adquirida por los especialistas, sino también porque los pacientes activos requieren citas de revisión, recuperación y acompañamiento del especialista [4]. Ante el alto volumen de pacientes, surge la necesidad de priorizar las citas de valoración, ya que los pacientes nuevos requieren tiempos de espera demasiado extensos para obtener una evaluación médica. Siendo la cita de valoración un requisito para la realización de una cirugía estética, se hace necesario identificar o establecer algunos criterios mínimos para la asignación prioritaria de las mismas.

Para realizar la priorización de pacientes, se tiene la información disponible en la historia médica, la cual contiene datos personales, antecedentes patológicos, medicamentos, antecedentes familiares, hábitos y evaluación física. En total, en la historia se almacenan alrededor de 85 variables, se requiere entonces, evaluar dicha información con el objetivo de identificar los factores que influyen en la decisión de realizarse una cirugía estética. La solución permitirá a un consultorio estético, asignar las citas de valoración priorizando aquellos pacientes que cumplen con las características identificadas. Las citas de valoración comúnmente son asignadas telefónicamente, lo que implica que la persona responsable de este proceso debe obtener por el mismo medio la información relacionada con los factores identificados para así realizar el proceso de priorización.

Para ejecutar la evaluación de las 85 variables, se recurre al área de Ciencia de Datos como un estudio interdisciplinar que involucra la aplicación de técnicas matemáticas, estadísticas, visualización de datos, aplicación de algoritmos computacionales para minería de datos y procesamiento en la nube. En la actualidad, la ciencia de datos se ha explorado debido al uso de “*Big Data*” y “*Data Analytics*”. Es en la analítica donde se espera resolver el problema de interés presentado en este

artículo, aplicando técnicas de minería de datos que permitan extraer información relevante de las historias clínicas de pacientes de cirugía estética.

Las técnicas de minería de datos se dividen en dos grandes ramas: predicción y descripción [5]. La predicción permite el pronóstico de eventos futuros sobre una variable objetivo a través de la aplicación de algoritmos supervisados como redes neuronales, árboles de decisión, máquinas de soporte vectorial, bayesianos, k vecinos más cercanos, entre otros. Por otro lado, la descripción permite la exploración de un conjunto de datos por medio de algoritmos no supervisados como k-means, cobweb, expectation-maximization, reglas de asociación, entre otros.

En el área de la salud, la minería de datos ha contribuido a la obtención de resultados predictivos y descriptivos sobre el padecimiento de enfermedades de alto impacto como cáncer, diabetes, leucemia, entre otras. En el área predictiva se encuentra el trabajo realizado por [6] para predecir la eficacia de la braquiterapia en el tratamiento de cáncer de próstata, ejecutando algoritmos supervisados como árboles de decisión. Adicionalmente se generaron reglas de asociación según las variables implicadas en la predicción y se propone ejecutar más adelante algoritmos bayesianos para contrastar o evaluar los resultados anteriormente obtenidos. Otro aporte al área predictiva se encuentra en [7], donde se realiza un estudio de patrones de supervivencia en mujeres con cáncer de cuello uterino mediante técnicas de árboles de decisión. También se hallaron reglas de clasificación obtenidas mediante la herramienta WEKA.

En el área descriptiva también se pueden encontrar aplicaciones de la minería de datos de la salud. En el trabajo realizado por [8], se encontraron las características comunes de pacientes diabéticos mediante una red neuronal de Kohonen, encontrando 3 perfiles diferentes. En la tesis escrita por [9], se realizó un estudio de minería de datos para conocer la comorbilidad asociada a la diabetes mellitus tipo 2 (DM2) en Colombia mediante dos algoritmos de agrupación: el algoritmo jerárquico (CHAMELEON) y el algoritmo particional k-means. El modelo estaba enfocado en dos partes; primero identificar el diagnóstico de comorbilidad y luego el tipo de diagnóstico. Fueron usadas finalmente redes bayesianas y árboles de decisión con un desempeño bueno y facilitaron de manera gráfica la

identificación de las características de los pacientes que presentaban este diagnóstico.

En la revisión literaria se pueden encontrar investigaciones en minería de datos de la salud que han logrado el reconocimiento de factores que influyen en el futuro diagnóstico de algunas enfermedades. En el estudio elaborado por [10], se reconoce la importancia de la minería de datos y su aplicación para la detección y el diagnóstico de pacientes con cáncer, enfocándose específicamente en el cáncer de mama, el uso de la mamografía y su interpretación. Se aplicaron diferentes métodos de forma independiente como PCA (*“Principal Components Analysis”*), árboles de decisión y reglas de asociación para encontrar relaciones no identificadas a primera vista. Otro análisis de factores fue realizado por [11], donde se utilizó la metodología CRISP-DM 1.0 para aplicar minería de datos a encuestas realizadas a pacientes diabéticos o con riesgo de desarrollarla. Se logró identificar y describir las características de los datos a emplear para clasificar pacientes de una localidad específica en Cuba y se identificaron atributos relevantes a la investigación. Otro reporte de análisis de factores se encuentra en [12], donde se aplicó minería de datos para identificar los factores de riesgo asociados a la muerte fetal de manera exitosa mediante algoritmos de clasificación y clustering obteniendo como resultado, las variables que más estaban correlacionadas con la muerte fetal.

En estos trabajos se reporta la aplicación tanto de técnicas predictivas como descriptivas para identificar los factores más relevantes que influyen en diferentes diagnósticos. A diferencia de estos trabajos, donde se aplican técnicas analíticas de forma independiente, en este artículo se propone un modelo analítico basado en un sistema de votación, donde se integran los resultados de varias técnicas. El modelo analítico propuesto se aplica en el análisis de factores que influyen en la realización de cirugías estéticas. La organización del artículo es el siguiente. En la sección 2 se describe la metodología, el modelo analítico propuesto y los datos analizados. En la sección 3 se presentan los resultados obtenidos y su respectivo análisis se presenta en la sección 4. Finalmente, en la sección 5 se presentan las conclusiones y trabajos futuros.

## 2. MATERIALES Y METODOS

A continuación, se describe la metodología de minería de datos utilizada como guía para este

estudio, el modelo analítico propuesto y la descripción de los datos analizados.

### 2.1 Metodología de Minería de Datos

En la literatura se pueden encontrar diversas metodologías de minería de datos como CRISP-DM, SEMMA, KDD, etc. Entre ellas se destaca CRISP-DM como una metodología completa y detallada del proceso de minería [13] [14] [15]. CRISP-DM plantea el desarrollo consecutivo de 6 etapas [16], llamadas comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue.

En la fase de comprensión del negocio se realizan las entrevistas necesarias con el especialista correspondiente; en este caso, el médico especialista en Cirugía Plástica, John Jairo Valencia Quintero, quien ejerce su profesión en la Clínica Las Américas en la ciudad de Medellín, Colombia. Entre los aspectos éticos del estudio, se aclara que durante el proceso de recolección de datos se respetó la Ley 23 Colombiana de 1981 donde se establecen de manera clara y específica los aspectos comportamentales y actitudinales del médico en la relación con el paciente y la conducta decorosa, respetuosa y digna, necesarias para hacer del proceso médico, un acto humanista

En la fase de comprensión de los datos se revisan las bases de datos entregados por el especialista, en este caso, correspondiente a historias clínicas y se procede al estudio de cada una de las variables con el fin de conocer el tipo de variable y sus valores posibles. Se identifican el número de variables y de registros. La fase de preparación de los datos plantea el desarrollo de procesos de selección de datos, limpieza y transformaciones que permitan obtener datos de calidad para la minería. Al igual que la fase anterior, es una etapa cuantitativa con un número de variables reducidas según la calidad de los datos. La fase de modelamiento consiste en primera instancia en seleccionar las técnicas analíticas adecuadas para el análisis y los datos. Se considera una etapa cuantitativa porque finaliza con los resultados arrojados por cada modelo. La fase de evaluación mide mediante métricas la calidad de los modelos desarrollados en la etapa anterior. Esta fase se caracteriza por ser cuantitativa ya que se expondrá de manera asertiva la justificación sobre la técnica seleccionada. Finalmente, en la fase de despliegue se analizan los resultados del proceso de minería y se establecen razones que lleven a una toma de

decisiones acertadas. En esta fase se justifican los resultados encontrados y se realizan sugerencias al especialista para la toma de decisiones, esto, mediante un informe cualitativo.

## 2.2 Modelo Analítico: Sistema de Votación

Para realizar el estudio de los factores que influyen en la realización de cirugías estéticas, en este artículo se propone un sistema de votación de métodos analíticos. Este tipo de sistemas han sido comúnmente utilizados en la analítica predictiva bajo el lema “dos expertos son más exactos que uno”. El sistema de votación propuesto se presenta en la figura 1, con la aplicación de las técnicas: análisis de correlaciones, análisis de componentes principales (PCA), árbol de decisión, regresión logística y reglas de asociación.

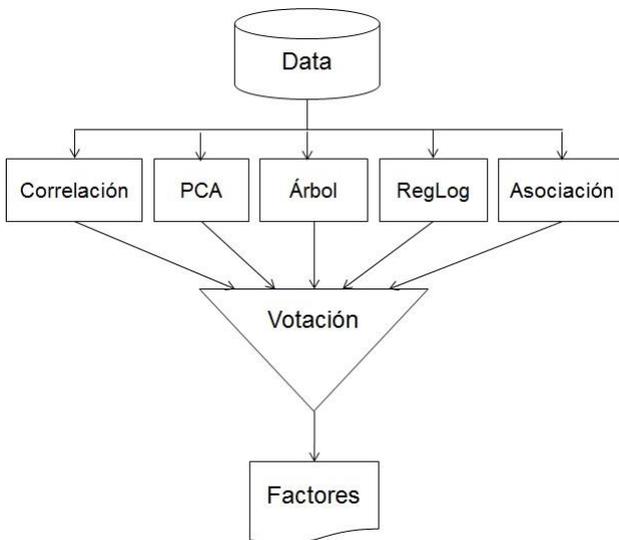


Fig 1. Sistema de votación propuesto.

Todas las técnicas aplicadas analizan el mismo conjunto de datos con el fin de buscar las relaciones existentes entre las variables de la historia clínica de los pacientes y la decisión de realizarse la cirugía estética. De la aplicación de cada técnica se obtiene una lista de los factores que más influyen para realizarse la cirugía. Para tomar la decisión final, se realiza una votación de los factores más seleccionados por las técnicas aplicadas.

## 2.3 Datos Analizados

Para realizar este estudio, se analizaron 2448 historias clínicas pertenecientes a los pacientes que han asistido a cita de valoración con el Cirujano Plástico John Jairo Valencia Quintero de la Clínica

Las Américas en Medellín – Colombia. Después de eliminar campos personales irrelevantes para el estudio de minería de datos se seleccionaron 57 variables agrupadas en: datos personales presentados en la tabla 1, antecedentes patológicos presentados en la tabla 2, medicamentos presentados en la tabla 3, antecedentes ginecológicos presentados en la tabla 4, antecedentes familiares presentados en la tabla 5, hábitos presentados en la tabla 6 y evaluación física presentada en la tabla 7.

Tabla 1. Variables con información personal.

| VARIABLE          | TIPO DE VARIABLE |
|-------------------|------------------|
| Sexo              | Catórica         |
| Edad Cirugía      | Catórica         |
| Estado Civil      | Catórica         |
| Ocupación         | Catórica         |
| Cirugía Deseada 1 | Catórica         |
| Cirugía Deseada 2 | Catórica         |

Tabla 2. Variables de antecedentes patológicos.

| VARIABLE              | TIPO DE VARIABLE |
|-----------------------|------------------|
| Tiroides              | Catórica         |
| Corazón               | Catórica         |
| Sida HIV              | Catórica         |
| Riñones               | Catórica         |
| Vesícula              | Catórica         |
| Artritis              | Catórica         |
| Problemas Coagulación | Catórica         |
| Presión Arterial      | Catórica         |
| Pulmones              | Catórica         |
| Sistema Nervioso      | Catórica         |
| Estomago              | Catórica         |
| Diabetes              | Catórica         |
| Asma                  | Catórica         |
| Lupus                 | Catórica         |
| Cáncer                | Catórica         |
| Fibromialgia          | Catórica         |
| Hígado                | Catórica         |
| Escleroderma          | Catórica         |
| Otras                 | Catórica         |

## 3. RESULTADOS

La experimentación fue realizada con la herramienta de minería de datos WEKA versión 3.8. WEKA (“Waikato Environment for Knowledge Analysis”) ha sido diseñada por un grupo de desarrolladores de la universidad de Waikato en Nueva Zelanda y es considerada una de las herramientas más usadas para minería de datos [17]. A continuación, se describen los resultados

obtenidos al aplicar las técnicas del sistema de votación.

**Tabla 3.** Variables de consumo de medicamentos.

| VARIABLE         | TIPO DE VARIABLE |
|------------------|------------------|
| Aspirina         | Categórica       |
| Ibuprofeno       | Categórica       |
| Para Artritis    | Categórica       |
| Para Dormir      | Categórica       |
| Para la Tos      | Categórica       |
| Para Tiroides    | Categórica       |
| Hormonas         | Categórica       |
| Tretinoína       | Categórica       |
| Antibióticos     | Categórica       |
| Para la Presión  | Categórica       |
| Para la Diabetes | Categórica       |
| Diuréticos       | Categórica       |
| Fenobarbital     | Categórica       |
| Digitálicos      | Categórica       |
| Cortisona        | Categórica       |
| Para Perder peso | Categórica       |

**Tabla 4.** Variables de antecedentes ginecológicos.

| VARIABLE                | TIPO DE VARIABLE |
|-------------------------|------------------|
| Edad Menstruacion       | Categórica       |
| Método de Planificacion | Categórica       |

**Tabla 5.** Variables de antecedentes familiares.

| VARIABLE                  | TIPO DE VARIABLE |
|---------------------------|------------------|
| Hipertensión Arterial     | Categórica       |
| Enfermedad Coronaria      | Categórica       |
| Artritis                  | Categórica       |
| Trastornos de Coagulación | Categórica       |
| Asma                      | Categórica       |
| Diabetes                  | Categórica       |
| Cáncer de mama            | Categórica       |
| Otros antecedentes        | Categórica       |

**Tabla 6.** Variables de hábitos.

| VARIABLE             | TIPO DE VARIABLE |
|----------------------|------------------|
| Cigarrillo           | Categórica       |
| Licor                | Categórica       |
| Café                 | Categórica       |
| Píldoras para dormir | Categórica       |
| Otros hábitos        | Categórica       |

**Tabla 7.** Variables de evaluación física.

| VARIABLE                   | TIPO DE VARIABLE |
|----------------------------|------------------|
| IMC (índice masa corporal) | Numérico         |

### 3.1 Selección de Factores con Análisis de Correlaciones

De acuerdo a los resultados arrojados por el análisis de correlaciones, las 10 variables valoradas en la tabla 8, juegan un papel importante dentro del modelo, ya que sus correlaciones con la variable objetivo son las más altas.

**Tabla 8.** Ranking con valor de correlaciones.

| RANKED | VARIABLE                       | VALOR   |
|--------|--------------------------------|---------|
| 1      | cirugía deseada 2              | 0.82628 |
| 2      | antecedente_familiar_hiperten. | 0.14981 |
| 3      | habito_licor                   | 0.10978 |
| 4      | habito_cafe                    | 0.10828 |
| 5      | IMC                            | 0.08274 |
| 6      | antecedente_familiar_diabetes  | 0.07954 |
| 7      | antecedente_familiar_otros     | 0.07823 |
| 8      | habito_cigarrillo              | 0.07627 |
| 9      | planificación                  | 0.07005 |
| 10     | cirugía deseada 1              | 0.06545 |

Según el análisis de correlaciones, la variable "Cirugía deseada 2" tiene la mayor incidencia sobre la realización de cirugías estéticas, con una correlación de 0,82628 sobre la variable estudiada, queriendo decir esto que, la segunda cirugía que el paciente manifiesta que desea realizarse en la consulta de valoración, es determinante para saber si realmente se realizará o no una cirugía. Es importante resaltar que también la variable "Cirugía deseada 1" hace parte de las 10 correlaciones más relevantes. Dentro del ranking de las 10 correlaciones más altas también se encontró que variables referentes a los antecedentes familiares de los pacientes influyen sobre la variable estudiada como antecedentes de Hipertensión Arterial, Diabetes y otros. Hábitos como el de consumir licor, consumir café y fumar fueron también relevantes. La influencia de la variable "IMC" (Índice de Masa Corporal) es también indiscutible ya que, por lo general, muchos pacientes que desean o se realizan liposucción y/o abdominoplastia presentan un índice de masa corporal alto. Por último, en el ranking la variable "Planificación" presentó una influencia sobre la variable estudiada.

### 3.2 Selección de Factores con PCA

La aplicación de la técnica de análisis de componentes principales (PCA – "Principal

Components Analysis”) arroja como resultado el ranking de factores de la tabla 9.

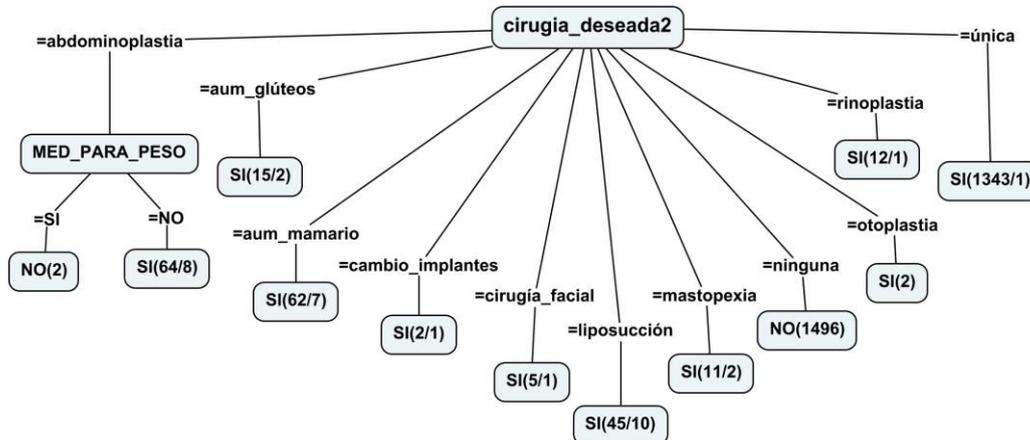
**Tabla 9.** Ranking de variables con PCA.

| RANKED | VARIABLE                         |
|--------|----------------------------------|
| 1      | IMC                              |
| 2      | cirugía deseada1=cirugía_facial  |
| 3      | cirugía deseada1=aum_mamario     |
| 4      | cirugía deseada1=bichectomía     |
| 5      | cirugía deseada1=braquioplastia  |
| 6      | cirugía deseada1=aum_implantes   |
| 7      | cirugía deseada1=aum_glúteos     |
| 8      | cirugía deseada1=abdominoplastia |
| 9      | ocupacion=servicio_doméstico     |
| 10     | ocupacion=servidor_público       |

En los resultados se encontró que la variable más importante en este método de análisis, es el “IMC” (Índice de Masa Corporal). Las siguientes 7 variables de mayor relevancia pertenecen a la primera cirugía que el paciente manifiesta que desea realizarse; algunas categorías son: cirugía facial, aumento mamario, bichectomía, braquioplastia, aumento de implantes, aumento de glúteos y abdominoplastia. Las últimas dos variables hacen referencia a ocupaciones de los pacientes; tales como: servicio doméstico y servidor público.

### 3.3 Selección de Factores con Árboles de Decisión

Al aplicar la técnica predictiva de árbol de decisión, se logra una exactitud del 98.7% en la tarea de predecir los pacientes que realizarán una cirugía estética. El árbol de aprendizaje generado se presenta en la figura 2.



**Fig 2.** Árbol de Decisiones.

En el árbol se puede observar que la variable más relevante para realizar la predicción es “cirugía deseada 2”. Los pacientes que NO se realizan un procedimiento quirúrgico son aquellos que manifiestan no desear una segunda cirugía y aquellos que desean una abdominoplastia y sí toman medicamento para perder peso. Por otro lado, los pacientes que SI se realizan el procedimiento quirúrgico son aquellos que como segunda cirugía manifiestan que desean una abdominoplastia y no toman medicamento para perder peso, un aumento de glúteos, un aumento

mamario, cambios de implantes, cirugía facial, liposucción, mastopexia, otoplastia, rinoplastia.

### 3.4 Selección de Factores con Regresión Logística

La regresión logística encuentra una ecuación de ajuste para predecir los pacientes que se realizarán una cirugía estética y los que no se realizarán el procedimiento. Al aplicar este método se logra una exactitud del 98.5% para predecir la realización del procedimiento. La tabla 10 presenta un resumen de las variables que influyen positivamente en la decisión de realizarse una cirugía estética según

las ecuaciones de regresión encontradas con este método.

Con las ecuaciones de regresión obtenidas en este experimento continúan siendo relevantes las cirugías estéticas que el paciente manifiesta que desea realizarse; el no consumo de algunos medicamentos; la no presentación de algunos antecedentes familiares y la ausencia de hábitos considerados como dañinos. Se presentan también algunas ocupaciones y estados civiles relevantes en cada una de las decisiones.

**Tabla 10.** Variables con afectación positiva en las ecuaciones de regresión.

| SI SE REALIZA LA CIRUGÍA   | NO SE REALIZA LA CIRUGÍA                                      |
|--|---|
| Estado civil= divorciado, soltero                                    | Estado civil= viudo, unión libre                              |
| Ocupación=empleado, docente, ingeniero                               | Ocupación=oficios varios, estudiante, independiente, operario |
| Cirugía deseada 1=mastopexia, aumento implantes, cambio de implantes | Cirugía deseada 1= liposucción                                |
| Cirugía deseada 2=otoplastia, rinoplastia                            | Cirugía deseada 2 = ninguna                                   |
| NO consumo de medicina, excepto para la presión                      | Sí consume medicinas  |
| Planificación= dispositivo, inyectable                               | Planificación= oral, ninguno                                  |
| Sin antecedentes familiares coronarios o diabetes                    | Con antecedentes familiares                                   |
| No tiene hábito de cigarrillo  |   |
| Sí tiene hábito de pastillas para dormir y otros                     |   |

### 3.5 Selección de Factores con Reglas de Asociación

Al aplicar el algoritmo A priori para crear reglas de asociación, se encuentran los siguientes resultados presentados en la tabla 11.

**Tabla 11.** Reglas de Asociación y la confianza de cada una.

| REGLA   | CONFIANZA |
|---|-----------|
| cirugia_deseada 2=ninguna<br>==> operado=no   | 100%      |
| cirugia_deseada2=ninguna<br>antecedente_familiar_coronaria=no<br>antecedente_familiar_artritis=no<br>habito_cafe=no habito_pastillas_dormir=no<br>habito_otros=no<br>==> operado=no   | 100%      |
| cirugia_deseada 2=única<br>IMC= (24.5-35)<br>==> operado=si   | 100%      |
| edad=juventud<br>cirugia_deseada2=única<br>antecedente_familiar_coronaria=no<br>antecedente_familiar_asma=no<br>antecedente_familiar_cancer=no<br>habito_pastillas_dormir=no<br>habito_otros=no<br>==> operado=si   | 100%      |
| cirugia_deseada2=única<br>antecedente_familiar_hipertenso=no<br>antecedente_familiar_coronaria=no<br>antecedente_familiar_artritis=no<br>antecedente_familiar_coagulacion=no<br>antecedente_familiar_diabetes=no<br>habito_cigarrillo=no<br>habito_licor=no habito_pastillas_dormir=no<br>habito_otros=no<br>==> operado=si | 100%      |
| medicina_para_artritis=no<br>medicina_para_tiroides=no<br>medicina_tretinoina=no<br>medicina_antibioticos=no<br>medicina_para_diabetes=no<br>medicina_fenobarbital=no<br>medicina_cortisona=no<br>==> operado=SI  | 63%       |
| medicina_para_artritis=no<br>medicina_para_tiroides=no<br>medicina_hormonas=no<br>medicina_tretinoina=no<br>medicina_antibioticos=no<br>medicina_para_diabetes=no<br>medicina_fenobarbital=no<br>medicina_digitalicos=no<br>==> operado=SI  | 63%       |

medicina\_para\_artritis=no  
 medicina\_para\_tiroides=no  
 medicina\_hormonas=no  
 medicina\_antibioticos=no  
 medicina\_para\_diabetes=no 63%  
 medicina\_fenobarbital=no  
 medicina\_cortisona=no  
 ==> operado=SI

medicina\_aspirina=no  
 medicina\_para\_tiroides=no  
 medicina\_hormonas=no  
 medicina\_antibioticos=no  
 medicina\_para\_presion=no 62%  
 medicina\_para\_diabetes=no  
 medicina\_diureticos=no  
 medicina\_fenobarbital=no  
 medicina\_digitalicos=no  
 medicina\_cortisona=no  
 medicina\_para\_peso=no  
 ==> operado=SI

Las dos primeras reglas generadas resaltan la importancia del no deseo sobre una segunda cirugía para que el paciente finalmente decida no operarse con una confianza del 100%, es decir que siempre se cumple esta regla. En cuanto a las reglas restantes, es relevante para la realización de un procedimiento estético tener un paciente sano, que no consume medicamentos, no tiene antecedentes familiares ni tiene malos hábitos.

#### 4. DISCUSIÓN

Para dar respuesta al objetivo planteado en este artículo, se procede a realizar la votación de las técnicas aplicadas para encontrar los factores que influyen en la realización de cirugías estéticas. La votación se presenta en la figura 3.

De la votación se obtienen los siguientes factores como los más relevantes: las cirugías deseadas en primera y segunda opción, el índice de masa corporal, hábito de cigarrillo y antecedentes familiares de diabetes. Los pacientes que SI se realizan un procedimiento quirúrgico son aquellos que:

- Manifiestan que desean una segunda cirugía
- La primera cirugía deseada es facial, aumento mamario, bichectomía, aumento de implantes, aumento de glúteos y abdominoplastia.
- El índice de masa corporal es normal
- No tienen hábitos de cigarrillo
- No tiene antecedentes familiares de diabetes

| VARIABLE                                  | EXPERIMENTOS DE MINERÍA DE DATOS |        |               |               |         |            |
|---|----------------------------------|--------|---------------|---------------|---------|------------|
|   | 1. Corre/                        | 2. PCA | 3. Árb de Dec | 4. Regr Logis | 5. Asoc | Nro. VECES |
| sexo                                      |                                  |        |               |               | X       | 1          |
| edad                                      |                                  |        |               | X             | X       | 2          |
| estado_civil                              |                                  |        |               | X             |         | 1          |
| ocupacion                                 |                                  | X      |               | X             |         | 2          |
| cirugia_deseada1                          | X                                | X      |               | X             |         | 3          |
| cirugia_deseada2                          | X                                |        | X             | X             | X       | 4          |
| medicina_aspirina                         |                                  |        |               |               | X       | 1          |
| medicina_ibuprofeno                       |                                  |        |               |               | X       | 1          |
| medicina_para_artritis                    |                                  |        |               | X             | X       | 2          |
| medicina_para_dormir                      |                                  |        |               |               | X       | 1          |
| medicina_para_tos                         |                                  |        |               | X             | X       | 2          |
| medicina_para_tiroides                    |                                  |        |               | X             | X       | 2          |
| medicina_hormonas                         |                                  |        |               |               | X       | 1          |
| medicina_tretinoina                       |                                  |        |               |               | X       | 1          |
| medicina_antibioticos                     |                                  |        |               |               | X       | 1          |
| medicina_para_presion                     |                                  |        |               | X             | X       | 2          |
| medicina_para_presion                     |                                  |        |               |               | X       | 1          |
| medicina_diureticos                       |                                  |        |               |               | X       | 1          |
| medicina_fenobarbital                     |                                  |        |               |               | X       | 1          |
| medicina_digitalicos                      |                                  |        |               |               | X       | 1          |
| medicina_cortisona                        |                                  |        |               |               | X       | 1          |
| medicina_para_peso                        |                                  |        |               | X             | X       | 2          |
| planificacion                             | X                                |        |               | X             |         | 2          |
| antecedente_familiar_hipertension         | X                                |        |               |               | X       | 2          |
| antecedente_familiar_enfermedad_coronaria |                                  |        |               | X             | X       | 2          |
| antecedente_familiar_artritis             |                                  |        |               | X             | X       | 2          |
| antecedente_familiar_coagulacion          |                                  |        |               |               | X       | 1          |
| antecedente_familiar_asma                 |                                  |        |               | X             | X       | 2          |
| antecedente_familiar_diabetes             | X                                |        |               | X             | X       | 3          |
| antecedente_familiar_cancer               |                                  |        |               |               | X       | 1          |
| antecedente_familiar_otros                | X                                |        |               |               | X       | 2          |
| habito_cigarrillo                         | X                                |        |               | X             | X       | 3          |
| habito_licor                              | X                                |        |               |               | X       | 2          |
| habito_cafe                               | X                                |        |               |               | X       | 2          |
| habito_pastillas_dormir                   |                                  |        |               | X             | X       | 2          |
| habito_otros                              |                                  |        |               | X             | X       | 2          |
| IMC                                       | X                                | X      |               |               | X       | 3          |

Fig 3. Resultado del sistema de votación.

**Factor 1 - Segunda cirugía deseada:** Se concluye que este factor es el más relevante en cuanto a la decisión de realizarse una cirugía estética. Muchos pacientes manifiestan su deseo de realizarse dos procedimientos en una sólo cirugía, por disminuir costos (gastos de quirófano, anestesia, instrumentación y honorarios médicos).

**Factor 2: Primera cirugía deseada:** Se determina importante entonces conocer qué tipo de cirugía motiva al paciente a asistir a una cita de valoración; pues, según el análisis realizado mediante minería de datos, las citas de valoración que finalmente terminan en una cirugía son las cirugías faciales, el aumento mamario, la bichectomía, el aumento de implantes, el aumento de glúteos y la abdominoplastia.

**Factor 3 - Índice de Masa Corporal:** Este factor es relevante para el área de la salud ya que, para que un paciente médicamente pueda realizarse una cirugía estética debe tener un índice de masa corporal adecuado; es decir, no deberá presentar

obesidad ni deberá estar por debajo del peso correspondiente a su talla; esto, por los riesgos que generan ambas condiciones para un procedimiento estético.

**Factor 4 - Hábitos de Cigarrillo** Es un factor importante para la realización de una cirugía estética, no tener el hábito de consumir cigarrillo. De hecho, el cirujano plástico John Jairo Valencia, exige a los pacientes que tienen el hábito de fumar, abandonarlo por lo menos dos meses antes del procedimiento estético ya que fumar cigarrillo afecta el proceso de oxigenación de los tejidos y por ende, el proceso de cicatrización generando deformidades a nivel cutáneo y también aumentando el riesgo de contraer infecciones, además, éste hábito genera problemas a nivel sistémico, es decir, a nivel pulmonar, cardíaco y cerebral.

**Factor 5 - Antecedentes Familiares de Diabetes:**

En general, los antecedentes familiares pueden ser importantes ya que aumentan el riesgo de que el paciente sufra una enfermedad específica. Llama la atención precisamente la aparición de éste antecedente ya que, con el paciente diabético se aumentan los riesgos de complicaciones en cirugía porque son pacientes que presentan menor aporte de nutrientes y de oxígeno a los tejidos en reparación. Además, este tipo de paciente puede tener daños en otros órganos internos asociados a su enfermedad. Es entonces importante estudiar qué importancia tiene para la realización de una cirugía estética que un paciente no presente antecedentes familiares de diabetes.

## 5. CONCLUSIONES Y TRABAJOS FUTUROS

En este artículo se desarrolló un estudio de los factores que influyen en la realización de cirugías estéticas mediante un modelo analítico, el cual fue aplicado en un conjunto de historias médicas de un consultorio estético ubicado la torre médica de la clínica Las Américas en la ciudad de Medellín, Colombia. El modelo analítico propuesto, el cual está basado en un sistema de votación, permite realizar un análisis de factores incluyendo los resultados obtenidos por varios métodos: análisis de correlaciones, análisis de componentes principales, árboles de decisión, regresión logística y reglas de asociación.

Sobre el modelo analítico se tiene que los métodos de correlaciones y análisis de componentes principales permiten determinar la medida de influencia de las variables sobre la decisión de realizarse la cirugía. Mientras que los métodos de árboles de decisión, regresión logística y a priori permiten determinar cómo es la influencia de las variables.

Con los resultados se encontró que los factores más relevantes para que un paciente se realice una cirugía estética en el consultorio del Doctor John Jairo Valencia Quintero son las cirugías que el paciente manifiesta desear realizarse en primer y segunda instancia. Otro factor corresponde a el índice de masa corporal, que, aunque en la mayoría de ocasiones es alto para quienes desean realizarse una cirugía estética corporal, no debe exceder los límites establecidos éticamente. Como tercer factor, se encontró que el hábito del paciente que consume cigarrillo influye en gran medida, esto por las implicaciones que tiene dicho consumo en el proceso de cicatrización. Por último, el factor de antecedentes familiares de diabetes, el cual es estudiado como un nuevo hallazgo y donde se sugirió realizar estudios que determinen el nivel de influencia entre dicho antecedente y la realización de una cirugía estética.

Para finalizar, se plantean algunas sugerencias de trabajo futuro en la creación de modelos analíticos que favorezcan el funcionamiento del consultorio como: 1) realizar un perfilamiento de pacientes, es decir, mediante algoritmo clustering, definir por agrupación, características de los pacientes del cirujano plástico John Jairo Valencia. 2) realizar un modelo predictivo de complicaciones quirúrgicas.

Adicionalmente, se propone utilizar el sistema de votación propuesto para realizar el análisis de factores en otras áreas y adicionar nuevos métodos a la votación.

## 6. AGRADECIMIENTOS

Los autores expresan su agradecimiento al grupo de investigación GIDATI de la Universidad Pontificia Bolivariana y al consultorio del Cirujano Plástico John Jairo Valencia Quintero. Este documento es resultado de la tesis de maestría en TIC titulada "MINERÍA DE DATOS DE LA SALUD: ANÁLISIS DE LOS FACTORES QUE INFLUYEN EN LA REALIZACIÓN DE CIRUGÍAS ESTÉTICAS".

## 7. REFERENCIAS BIBLIOGRÁFICAS

- [1] Elliott A., Making the Cut: How Cosmetic Surgery is Transforming Our Lives, London, 2008
- [2] ASPS. Plastic Surgery Statistic. Disponible en: <https://d2wirczt3b6wjm.cloudfront.net/News/Statistics/2016/2016-plastic-surgery-statistics-report.pdf> [consultado el 25 de junio de 2017].
- [3] ISAPS. International Survey Aesthetic /Cosmetic Procedures Performed, 1-18, 2014
- [4] Satisfacción y Calidad Percibida en la Atención de Salud Hospitalaria: Ranking de Prestadoras – Informe Global. Chile, Superintendencia de Salud, 2013.
- [5] Kumari M, Sunila G. Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction, 2, 2011.
- [6] Reparaz, D., Merlino, H., Rancan, C., Rodríguez, D., Britos, P., & García Martínez, R., Determinación de la Eficacia de la braquiterapia en tratamiento de cáncer basada en minería de datos. Universidad de Buenos Aires, Facultad de Ingeniería, Buenos Aires, 2008.
- [7] Timarán, R., & Yépez, M. C., La minería de datos aplicada al descubrimiento de patrones de supervivencia en mujeres con cáncer invasivo de cuello uterino, Universidad y Salud 14(2); 117-129, 2012.
- [8] Jiménez, M. I., & Mateos, R., Perfiles de pacientes diabéticos que sufren reacciones adversas a medicamentos. Identificación y caracterización a través de minería de datos; An. Real Acad, Farm, 80(2), 274-321, 2013.
- [9] Franco, Á., M., Modelo para Análisis de Riesgo de la Diabetes Mellitus 2 usando Inteligencia de Negocios y Minería de Datos [Tesis de Maestría], Bogotá: Universidad Nacional de Colombia, 2014.
- [10] Rodríguez, F., & Vallejo, N. Aplicación de técnicas de Minería de Datos para el diagnóstico prematuro de Cáncer. Universidad. Madrid: Universidad Calos III, 2012.
- [11] Acosta, R., Rosete, A., & Rodríguez, A., Predicción de pacientes diabéticos. Preprocesado para Minería de Datos. Ciudad Universitaria José Antonio, La Habana: Ciudad Universitaria José Antonio Echeverría (CUJAE), Facultad de Ingeniería Informática, 2012.
- [12] Toscano de la Torre, B.A, Ponce Gallegos, J.C, Margain Fuentes, M.L, López Espinosa, R, Meza de Luna, M.A, en Aplicación de Minería de Datos para la Identificación de Factores de Riesgo Asociados a la Muerte Fetal, CiComp, 215-224, 2016.
- [13] Moine J. Metodologías para el descubrimiento de conocimiento en bases de datos: un estudio comparativo [Tesis de Maestría]. Argentina: Universidad Nacional de la Plata, 2013.
- [14] Torres D. Diseño y aplicación de una metodología para análisis de noticias policiales utilizando minería de textos [Tesis de Pregado]. Chile: Universidad de Chike, 2013.
- [15] Corrales D, Ledesma A, Peña A, Hoyos J, Figueroa A, Corrales J. A new dataset for coffee rust detection in Colombian crops base on classifiers. Revista S&T, 9-23, 2014.
- [16] IBM, Manual CRISP- DEM de IBM SPSS Modeler, EEUU, 2012.
- [17] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann, Witten IH. The WEKA Data Mining Software: An Update. SIGKDD Explorations, 10-18, 2009.