



**UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS**

**FACULTAD 7**

**TRABAJO DE DIPLOMA PARA OPTAR POR EL TÍTULO DE INGENIERO EN CIENCIAS  
INFORMÁTICAS**

**SISTEMA BASADO EN CASOS PARA PREDECIR LA OCURRENCIA DE REACCIONES ADVERSAS A  
MEDICAMENTOS EN LA CONSULTA MÉDICA**

**AUTORES**

Frank David Corona Prendes

Luis Enrique Trujillo Irarragorri

**TUTORES**

MSc. Annia Arencibia Morales

Ing. Yoiler J. Frómeta Moreno

**La Habana, junio de 2013**  
"Año 55 de la Revolución"

*“Toda persona que utiliza profesionalmente un medicamento tiene que preocuparse activamente por las consecuencias de su utilización y recordar que no hay rosa terapéutica sin espinas adversas”.*



### **DATOS DE CONTACTO**

MSc. Annia Arencibia Morales (aarencibia@uci.cu): graduada de Ingeniero en Ciencias Informáticas en la Universidad de las Ciencias Informáticas en el año 2007. Profesora de la Universidad de las Ciencias Informáticas, pertenece al Centro de Informática Médica (CESIM), actualmente posee la categoría docente de profesor Asistente. Máster en Informática Aplicada. Profesora de Metodología de la Investigación Científica e imparte Postgrado de Ingeniería de Software. Pertenece al Departamento de Sistemas de Apoyo a la Salud.

Ing. Yoiler J. Frómeta Moreno (yfrometa@uci.cu): graduado de Ingeniero en Ciencias Informáticas en la 1era graduación de la Universidad de las Ciencias Informáticas (UCI). Ha impartido clases de Metodología de la Investigación y Práctica Profesional entre otras. Actualmente posee la categoría docente de Instructor y es miembro del Departamento Sistema de Apoyo a la Salud en el Centro de Informática Médica (CESIM), donde ejerce como desarrollador del producto Synta que cuenta con cinco módulos en desarrollo.

## AGRADECIMIENTOS

**Frank David Corona Prendes:**

*Agradecer a la Revolución cubana por haberme permitido estudiar en esta institución y brindarme los medios para convertirme en un profesional que le pueda servir.*

*A mi familia, por haber estado ahí desde el principio, por su apoyo incondicional; siempre dándome sus consejos.*

*A mi madre, por sus regaños que tanto me enseñaron, por el cariño, la comprensión y la dedicación a escucharme, por hacerme tu confidente,*

*A mi padre, por esas conversaciones sobre la vida, siempre voy a recordar esa conversación que tuvimos antes de entrar en la universidad, por las charlas científicas, siempre enseñándome y alentándome a ser mejor.*

*Mi padrastro por ser como un padre; por cuidar de mi mamá y preocuparse tanto por nosotros y no pensarlo dos veces cuando tenía que salir a buscar algo para la casa, por los momentos que estuvimos escuchando tus cuentos de Etiopía y el Tanquista.*

*A mi hermano por ser mi aliento a seguir adelante, por ponerme el reto de ser tu guía, ¡tremenda tarea me pusiste brother!, por los tragos que nos dimos juntos, por las charlas e intercambio de experiencias, en fin por estar siempre ahí, por ser mi familia más cercana en la universidad.*

*A mis abuelos siempre tan suspicaces con su experiencia, la cual nunca dudaron de utilizarla para guiarnos ¡mis viejos, los quiero con la vida!*

*A mi novia quien nunca me ha fallado un año, un mes y 16 días, gracias por tus cariños, tu presencia y tu comprensión ¡te quiero cosita!*

*A mi compañero de tesis Luisito, mi hermano, gracias por tu amistad incondicional, por todo el tiempo que compartimos juntos rompiéndonos la cabeza. De ser mi amigo pasaste a ser un hermano.*

*A mis amistades de la universidad, mi gente de clases Alexis, Rey, Nane, Arlety, Nelson, Hansel mi primo perdido, Yunior, Andres, David, todas mis amistades de la 4 el Yuni, Olga, David el Dulce, el Under. A mis amistades que ya se graduaron, Anniel, Nairelis, Jose, en fin a todos aquellos que me enseñaron lo más bueno de tener amigos.*

*Mis tutores, Annia, gracias por pensar en mí para esta investigación por guiarme desde segundo año cuando entre al proyecto, por darnos tus charlas en aras de hacernos mejores, por las madrugadas que estuvimos en tu casa arreglando las cosas ¡ves el negro se hizo ingeniero!, Joaquín por la guía y los consejos, al profe Avelino; nunca podré agradecerle por todo lo que hizo por mí en primer año, Darien, Arieskien, Pompa, Mojena, Guillermo por todo el tiempo que me dedicaron cuando me atascaba con la programación ¡de verdad gracias!*

## **Luis Enrique Trujillo Irarragorri:**

*Le agradezco a mi mamá (la persona más tierna y amorosa que existe) por estar siempre conmigo ayudándome, apoyándome y obligándome siempre a tomar por el buen camino y siempre pensando en mi futuro y bienestar.*

*A mi papá por enseñarme tantas cosas de la vida, siendo siempre un gran amigo para mí.*

*A mi padrastro que siempre estuvo para mí cuando se me presentaba alguna dificultad o cuando necesitaba consejos.*

*A mi esposa que me enseñó lo que necesitaba con amor.*

*A mi abuela Angélica y mi abuelo Pepe (que no se encuentra entre nosotros) que me criaron y malcriaron desde pequeño. A mi abuela Alba y mi abuelo Felipe (que no se encuentra entre nosotros) que siempre me llevaron a pasear y les gustaba que me divirtiera.*

*A todos mis tíos y mis primos que me han ayudado y apoyado en todo lo que les ha sido posible.*

*A Dennier, Darien, Víctor, Dainier, Rogeidys, Reinaldo, Alejandro, Carlos Omar, Héctor, Antonio, Abraham, Lisandra y Michele por su amistad y estar conmigo en las buenas y en las malas. De forma general a todos los amigos de mi tierra y de la Universidad que me enseñaron lo hermoso de la amistad.*

*A mi compañero de tesis por lograr complementarnos tan bien para poder lograr cumplir con nuestra tarea.*

*A mis compañeros de apartamento: Jacke, Aniuska, Yanetsy, Lili y Felix por compartir en su compañía momentos inolvidables. A mis compañeros de grupo.*

*A Annia (tutora) por tenerme siempre corriendo, asustado y bajo presión. Por su especial dedicación creyendo que era su hijo. Por ser una buena tutora, profesora y amiga.*

## Agradecimientos

---

*A los profesores Arieskien, Pompa, Hugo, Yoiler (tutor), Mainoldis por todo el apoyo que me brindaron siempre que necesité de su colaboración. A todos los profesores que de alguna u otra forma contribuyeron en mi formación como ser humano y futuro profesional.*

## **DEDICATORIA**

*Frank David Corona Prendes:*

Dedico el trabajo a mis padres y mis hermanos. Mamá, papá aquí está su hijo como profesional, espero sea capaz de reconfortar el esfuerzo que pusieron en mí, hermano espero haber cumplido bien la tarea de guiarte y si yo pude, ¡tú también puedes!

*Luis Enrique Trujillo Irarragorri:*

Esta tesis se la dedico a mi mamá, a mi papá, a mi padrastro, a mis abuelos, a mi esposa, familia, amigos y profesores por formar una parte importante de mi vida. A mi bebé que está por nacer.

## RESUMEN

Las Reacciones Adversas a Medicamentos (RAM) pueden causar incapacidad temporal o permanente al paciente. Incluso tener un desenlace fatal, figurando entre las diez causas principales de defunción en todo el mundo. La ocurrencia de las RAM, presentan costos directos teniendo en cuenta: salario del personal, gasto de material y costo de los medicamentos, los cuales aumentan la cuantía al tratarla. Para predecir una posible RAM se realiza un Sistema Basado en Casos (SBC) que le permite al médico consultar qué RAM puede ocurrir en el paciente ante un nuevo tratamiento.

Para la creación del SBC, se utilizó el Triángulo de Fuller para calcular el peso de los rasgos objetivos y para la determinación de la incertidumbre, los estadísticos Chi-Cuadrado y V-Cramer. Al realizar el cálculo se separaron los rasgos predictores en asociados y no asociados, culminando con la creación del árbol de decisión. Posteriormente se aplicó la Función de Semejanza y Similitud, seleccionando de todos los casos el más óptimo.

El SBC se integró al Módulo de Reacciones Adversas a Medicamentos, del producto Synta, desarrollado por el Departamento de Sistemas de Apoyo a la Salud (SAS), del cual obtiene la información necesaria referente a las RAM; utilizando esta para la predicción. A la solución obtenida se le realizaron pruebas en aras de obtener el grado de certeza de las respuestas; los resultados obtenidos oscilan entre el 70% y 80%, reflejando un alto grado de certeza en la predicción de posibles RAM.

Palabras clave: *predicción, Reacciones Adversas a Medicamentos, Sistema Basado en Casos*



## ÍNDICE

Introducción .....	4
Capítulo 1: Fundamentación teórica de la investigación.....	11
1.1 Soluciones existentes a nivel internacional y nacional .....	11
1.1.1 Sistema Experto para el Diagnóstico de Enfermedades Infecciosas en la Sangre (MYCIN) .	11
1.1.2 Razonamiento Basado en casos en ciencias medicas sobre plataforma web.....	11
1.1.3 Sistema experto para el diagnóstico neurológico (NDS) .....	12
1.1.4 Sistema Inteligente para la Selección de Información .....	13
1.2 Sistemas Expertos.....	13
1.2.1 Estructura de un Sistema Experto.....	14
1.2.2 Clasificación de los Sistemas Expertos más usados en la salud.....	14
1.2.3 Tipos de búsqueda en los Sistemas Expertos .....	18
1.2.4 Incertidumbre en los SBC .....	19
1.3 Modelos de base de datos .....	20
1.3.1 Modelo Plana.....	20
1.3.2 Modelo Jerárquico .....	20
1.3.3 Modelo de datos en red .....	21
1.3.4 Modelo de datos relacional.....	22
1.4 Identificación de los rasgos predictores .....	22
1.4.1 Algoritmos existentes para el cálculo de los Testores Típicos.....	22
1.5 Herramientas y tecnologías a utilizar .....	24
1.5.1 Marcos de desarrollo .....	24
1.5.2 Symfony 1.4.....	25
1.5.3 HTML.....	25
1.5.4 CSS.....	25
1.5.5 JavaScript 1.1 .....	26
1.5.6 ExtJS 3.1 .....	26
1.5.7 Lenguaje de Programación.....	26
1.5.8 NetBeans 6.9.....	27

1.5.9	Servidor Web Apache 2.2 .....	27
1.5.10	Sistema gestor de base de datos MySQL 5.0 .....	27
1.5.11	EMS MySQL Manager 3.7 .....	27
Capítulo 2: Propuesta de solución del Sistema Basado en Casos .....		28
2.1	Propuesta de solución .....	28
2.2	Técnicas utilizadas en el desarrollo del SBC.....	29
2.3	Flujo del modelo a implementar .....	37
Capítulo 3: Validación del Sistema Basado en Casos.....		47
3.1	Interfaz de usuario .....	47
3.2	Principio de la parsimonia o Navaja de Ockham .....	49
3.3	Aplicación del principio de la parsimonia o Navaja de Ockham en la validación de la solución....	49
Conclusiones .....		54
Recomendaciones .....		55
Referencias bibliográficas .....		56
Bibliografía.....		60
Anexos.....		66

## Introducción

La salud humana es de extrema importancia en cualquier sociedad; para lograr el mejoramiento de la misma, los diplomáticos que se reunieron para crear las Naciones Unidas en 1945, abordaron diversos temas, entre los cuales se encontraba la posibilidad de establecer una organización mundial dedicada al cuidado de la salud de los pacientes. El 7 de abril de 1948 entró en vigor la Constitución de la Organización Mundial de la Salud (OMS), autoridad directiva y coordinadora de la acción sanitaria en el sistema de las Naciones Unidas, que tenía la responsabilidad de desempeñar una función de liderazgo en los asuntos sanitarios mundiales, configurar la agenda de las investigaciones en salud, establecer normas, articular opciones de políticas basadas en la evidencia, prestar apoyo técnico a los países y vigilar las tendencias sanitarias mundiales. (OMS, 2013)

Uno de los programas de la OMS, es el Programa Internacional de Farmacovigilancia, que en el año 1968 se pone en marcha, con la idea de aglutinar los datos existentes sobre las Reacciones Adversas a los Medicamentos (RAM). Este programa estuvo orientado a la toma de decisiones, que permitían mantener la relación beneficio-riesgo de los medicamentos en una situación favorable, o bien suspender su uso cuando esto no sea posible. Existen ochenta y seis países que participan en el programa coordinado por la OMS y su Centro Colaborador de Uppsala (ver Anexo 1). (OMS, 2012)

La RAM representan cualquier efecto perjudicial o indeseado, que ocurre tras la administración de un fármaco normalmente utilizado para la profilaxis<sup>1</sup>, diagnóstico o tratamiento, que modifica desfavorablemente el curso clínico del síntoma o de la enfermedad tratada, causando incomodidad o agravamiento general del paciente y que aparece después de la administración de dosis terapéuticas habituales (OMS, 2013). El Programa Internacional de Farmacovigilancia, desarrolló métodos para mantener el control de los fármacos comercializados y definió las RAM por impacto y causalidad. Estos métodos de Farmacovigilancia son:

*Sistema de Notificaciones Espontáneas*: se basa en la identificación y detección de las sospechas de RAM por parte de los profesionales de la salud en su práctica diaria, y el envío de esta información a un organismo que la centraliza. (ANMAT, 2009)

---

<sup>1</sup> Conjunto de medidas que se aplican para prevenir las enfermedades.

Sistemas de Farmacovigilancia Intensiva: se fundamentan en la recolección de datos en forma sistemática y detallada de todos los efectos perjudiciales, que pueden concebirse como inducidos por los medicamentos, en grupos bien definidos de la población y se dividen en dos grandes grupos (ANMAT, 2009):

- Sistemas centrados en el medicamento
- Sistemas centrados en el paciente

El más difundido de los métodos en el estudio de la Farmacovigilancia, es el Sistema de Notificaciones Espontáneas, también llamado Tarjeta Amarilla; surgido en España en 1982. La notificación sistemática de reacciones adversas y su análisis estadístico permanente, permiten generar señales de alerta sobre el comportamiento de los medicamentos en la población. El éxito o fracaso de cualquier actividad de Farmacovigilancia, depende de la notificación de las sospechas de RAM (Uppsala, et al., 2001). Los fármacos actuales, son capaces de modificar profundamente los procesos fisiológicos y bioquímicos normales de un paciente y su eficacia terapéutica está indisolublemente ligada a su capacidad de producir RAM.

Estudios recientes, en la Farmacovigilancia, revelan que el uso de los medicamentos y sus consecuencias, cuestan millones de dólares al año y figuran entre las diez causas principales de defunción en todo el mundo. Las RAM afectan a personas de todos los países. En algunos casos, los costos asociados a estas sobrepasan el precio de los medicamentos utilizados en los tratamientos de los pacientes. Al menos un 60% de RAM son evitables y sus causas (OMS, 2013) pueden ser las siguientes:

- Diagnóstico erróneo
- Prescripción del medicamento equivocado o de una dosis equivocada del medicamento correcto
- Trastornos médicos, genéticos o alérgicos subyacentes que pueden provocar una RAM
- Automedicación con medicamentos que requieren prescripción
- Incumplimiento del tratamiento prescrito
- Reacciones con otros medicamentos (incluidos los medicamentos tradicionales) y determinados alimentos

- Uso de medicamentos de calidad inferior a la norma, cuyos ingredientes y composición no se ajustan a las especificaciones científicas apropiadas, y que pueden resultar ineficaces y a menudo peligrosos
- Uso de medicamentos falsificados sin ingredientes activos o con ingredientes equivocados, que pueden ser peligrosos y hasta mortales

Anticiparse a quién puede sufrir una RAM, cuándo y qué medicamento la causa; es una dificultad, porque no existe un grupo de pacientes con similares características que indiquen la ocurrencia de una RAM. Las investigaciones realizadas no han identificado un predecesor válido de estos eventos. Actualmente las características de los pacientes no son útiles predictores de las RAM, debido a que no existe un grupo homogéneo de pacientes que padezcan de reacciones adversas. Para realizar un trabajo organizado, la OMS clasifica las RAM según su severidad y relación de causalidad (Laporte, y otros, 2007 pág. 148):

#### Según su severidad:

*Leve:* se consideran en este grupo las reacciones que no amenazan la vida del paciente, no provocan hospitalización, ni atención en servicios de urgencias, ni baja laboral.

*Moderada:* la reacción provoca hospitalización o atención en servicios de urgencias, o baja laboral o escolar, sin amenazar directamente la vida del paciente.

*Grave:* la reacción amenaza directamente la vida del paciente.

*Mortal:* provoca la muerte del paciente.

#### Relación de causalidad

*No relacionada:* no existe un intervalo adecuado entre el uso del medicamento y la reacción.

*Condiciona:* existe un intervalo adecuado entre el uso de medicamento y la reacción, es una reacción no conocida. La reacción se puede explicar por el cuadro clínico o por otro fármaco.

*Posible:* se considera cuando existe un intervalo adecuado entre el uso del medicamento y la reacción, es una reacción conocida. La reacción no se puede explicar por el cuadro clínico o por otro fármaco y no existe reaparición de la reacción tras la reexposición al fármaco.

*Probable:* existe un intervalo adecuado entre el uso del medicamento y la reacción, es conocida; no se puede explicar por el cuadro clínico o por otro fármaco y el paciente mejora al suspender la medicación.

*Definitiva:* existe un intervalo adecuado entre el uso de medicamento y la reacción; es conocida; no se puede explicar por el cuadro clínico o por otro fármaco; al suspender la medicación existe mejoría y reaparición tras la reexposición.

Cuba, es uno de los países que pone en práctica el Programa Internacional de Farmacovigilancia. El desarrollo del mismo, ha permitido detectar una elevada tasa de reacciones adversas por medicamentos (de siete mil a diez mil casos anuales), con más de seiscientas asociaciones de fármaco-reacción por millón de habitantes. A través de los años, con el uso del Programa Internacional de Farmacovigilancia, se ha perfeccionado la detección y el análisis de los efectos adversos, aumentando el registro de efectos adversos moderados y graves, que generalmente requieren atención de urgencia y/o ingreso hospitalario (ver Anexos 2 y 3). Este perfeccionamiento que se ha llevado a cabo continúa siendo insuficiente, debido a que todas estas notificaciones son registradas por el médico (o personal de salud), haciendo uso de una planilla distribuida por el Ministerio de Salud Pública (MINSAP) a todas las Unidades de Salud (US). (Farmacovigilancia, 2010)

La información mostrada en los Anexos 2 y 3 representa un aproximado de las RAM ocurridas en Cuba, debido a que existe una parte de la misma que no es válida para analizar, por no ser confiable. Los profesionales encargados de registrar los datos de las RAM, en la planilla distribuida por el MINSAP, no lo hacen de forma correcta, dejan información importante sin adjuntar y en muchos casos no se recogen los datos correspondientes a las RAM, por no poseer en las unidades de salud toda la información necesaria.

La ocurrencia de las RAM provoca gastos por partidas, tales como: el salario, gasto de material alimentos, material de curación, material de laboratorio y radiológico, instrumental médico, etc.; aumentando la cuantía de tratar una reacción adversa. El costo de RAM evitables ocupa un 36% del costo de ingreso por RAM. Estudios realizados por especialistas de Farmacovigilancia en los hospitales Joaquín Albarrán y Salvador Allende de La Habana, demuestran que de un total de ochenta y dos ingresos por RAM, veintiséis reacciones adversas podían haber sido prevenibles, representando un 31,7%, del total. (ver Anexos 4 y 5) (Farmacovigilancia, 2010)

Cada año son introducidos y retirados del Formulario Nacional de Medicamentos (FNM)<sup>2</sup>, nuevos medicamentos importados o producidos en Cuba. En el año 2009, debido al monitoreo activo que se realizó, se evidenció un aumento de los reportes de sospechas de RAM, donde se identificó, que más del 20% del presupuesto asignado a la salud fue empleado en complicaciones producidas por las RAM (ver Anexo 6). (Farmacovigilancia, 2010)

Estos gastos asociados a las Reacciones Adversas a Medicamentos también se evidencian a nivel internacional. En los hospitales de EE.UU, más de 770 000 pacientes son perjudicados o mueren cada año por esta causa, lo cual conlleva a un gasto de más de 5,6 millones de dólares al año por hospital (Quality, 2001). Otros estudios, estiman que un 9,7% de las RAM causan incapacidad permanente. Además existe un riesgo de muerte en pacientes que han sufrido RAM dos veces mayor que en uno que no la padeció. En investigaciones realizadas en el Hospital Brigham and Women's y el Hospital General de Massachusetts, se identificó que las RAM aumentan la estadía de los pacientes de cuatro a seis días, incrementando los gastos. (Quality, 2001)

Las hospitalizaciones por RAM se incrementan en personas que han vivido más tiempo, debido a que presentan un mayor número de enfermedades crónicas, aumentando de esta forma la cantidad de medicamentos consumidos. Los adultos de sesenta y cinco años de edad o mayores, que toman entre cinco y nueve medicamentos, representan el 40%, mientras que un 18% toman diez ó más medicamentos. Los adultos mayores tienen siete veces más probabilidad que las personas jóvenes a tener efectos adversos a fármacos que requieran hospitalización (ver Anexo 7). (Budnitz, et al., 2012)

Por todo lo anteriormente planteado se presenta como **problema a resolver**: ¿Cómo contribuir a la predicción de la ocurrencia de las Reacciones Adversas a Medicamentos en los pacientes, desde la consulta médica?

Este problema está enmarcado en el **objeto de estudio**: Sistemas Expertos en el campo de la salud, centrándose en el **campo de acción**: Sistema Basado en Casos para la predicción de la ocurrencia de Reacciones Adversas a Medicamentos en los pacientes, desde la consulta médica.

---

<sup>2</sup> Formulario Nacional de Medicamentos: listado de medicamentos (elaborado por el MINSAP) donde se registran los productos farmacéuticos que se utilizan para la atención a la salud cubana.

Para solucionar el problema planteado, se define como **objetivo general**: desarrollar un Sistema Basado en Casos que permita la predicción de la ocurrencia de las Reacciones Adversas a Medicamentos en los pacientes, desde la consulta médica.

Como tareas de la investigación se proponen las siguientes:

1. Analizar los Sistemas Expertos a nivel internacional y nacional utilizados en la predicción médica, estableciendo similitudes con la investigación en curso.
2. Analizar la estructura y composición de los Sistemas Expertos, seleccionando el más idóneo para la predicción de las Reacción Adversa a Medicamentos desde la consulta médica, teniendo en cuenta la información existente.
3. Asimilar las herramientas y tecnologías propuestas por el Departamento de Sistemas de Apoyo a la Salud, para solucionar el problema planteado.
4. Identificar los rasgos predictores y objetivos a utilizar en la Base de Casos y en la Máquina de Inferencia del Sistema Basado en Casos.
5. Crear una Base de Casos, donde se almacene la información referente a los registros de las Reacciones Adversas a Medicamentos, para ser utilizada en la Máquina de Inferencia.
6. Crear una Máquina de Inferencia que permita predecir la ocurrencia de las Reacciones Adversas a Medicamentos, teniendo en cuenta la información existente en la Base de Casos.
7. Realizar las interfaces de usuario del sistema, siguiendo las pautas definidas por el Departamento de Sistemas de Apoyo a la Salud, para el desarrollo del proyecto Synta.

Entre los resultados que se esperan obtener con el desarrollo de la presente investigación, se encuentran:

- Servirá como mecanismo de apoyo para la predicción de la ocurrencia de las Reacciones Adversas a Medicamentos en los pacientes, desde la consulta médica.
- Permitirá minimizar los costos asociados a las Reacciones Adversas a Medicamentos.

Para el desarrollo del trabajo se utilizaron métodos de investigación. A continuación se listan los mismos:



## **Métodos Teóricos**

*Análisis histórico-lógico:* se utilizó para realizar la valoración de Sistemas Expertos existentes a nivel internacional y nacional. Para conocer el funcionamiento y estructura de los Sistemas Expertos y haciendo un estudio y valoración de los Testores Típicos, identificando cuál de ellos es el más efectivo.

*Analítico-Sintético:* se puso en práctica para determinar las técnicas a utilizar en la máquina de inferencia.

*Inductivo-Deductivo:* se utilizó en la aplicación de casos de pruebas al sistema, llegando a conclusiones a partir de las respuestas proporcionadas por este.

## **Métodos Empíricos**

*Entrevistas:* se realizaron entrevistas estructuradas a los farmacoepidemiólogos, identificando la información necesaria para la creación del Sistema Basado en Casos.

El documento está estructurado en tres capítulos, siendo estos:

**Capítulo1: Fundamentación teórica de la investigación**, incluye un análisis y revisión del estado del arte, a nivel internacional y nacional, de Sistemas Expertos en el campo de la salud. Se hace una descripción de las tendencias, técnicas, tecnologías, herramientas y algoritmos usados para el desarrollo del Sistema Basado en Casos.

**Capítulo2: Propuesta de solución del Sistema Basado en Casos**, se explican las técnicas seleccionadas para la creación del Sistema Basado en Casos, brindando un ejemplo práctico sobre la utilización de las mismas.

**Capítulo 3: Validación del Sistema Basado en Casos**, se muestran los resultados de las pruebas internas realizadas, teniendo en cuenta la valoración del experto.

## Capítulo 1: Fundamentación teórica de la investigación

Para el campo de la salud se han desarrollado Sistemas Expertos, cuyo objetivo es el diagnóstico de enfermedades y recomendaciones de tratamientos, entre estos se puede mencionar:

### 1.1 Soluciones existentes a nivel internacional y nacional

#### 1.1.1 Sistema Experto para el Diagnóstico de Enfermedades Infecciosas en la Sangre (MYCIN)

Una de las primeras manifestaciones exitosas de Sistemas Expertos, se produjo precisamente en el campo de las Ciencias Médicas, en la década del setenta con el desarrollo del sistema MYCIN. Este sistema fue concebido para el diagnóstico de enfermedades infecciosas en la sangre. El mismo fue desarrollado por Edgar Short Liffe, en la Universidad de Stanford, escrito en lenguaje de programación List Processing (LISP) e inspirado en el Sistema Experto Dendral. (Santealla, 2005)

MYCIN es capaz de razonar el proceso seguido por el especialista para llegar a sus diagnósticos y de recetar medicaciones personalizadas a cada paciente (según su estatura y peso entre otras características). El sistema diagnostica la causa de la infección usando el conocimiento relativo a la infección de los microorganismos, a partir de historiales de pacientes, síntomas y los resultados de los test de laboratorio. MYCIN se basa en realizar los razonamientos en factores de certeza (Santealla, 2005)

MYCIN es un sistema robusto para determinar correctamente conclusiones cuando algunas evidencias son incompletas o incorrectas, puede explicar sus propios procesos de razonamiento y es uno de los primeros sistemas genuinamente amigables. Razona heurísticamente, interactuando con los humanos utilizando lenguaje natural y con descripciones simbólicas. Además de que puede contemplar hipótesis competitiva. En investigaciones realizadas por Stanford Medical School, se determinó que MYCIN tenía una tasa de aciertos de aproximadamente el 65%, lo cual mejoraba las estadísticas de la mayoría de los médicos no especializados en el diagnóstico de infecciones bacterianas para la década de los setenta. (Santealla, 2005)

#### 1.1.2 Razonamiento Basado en casos en ciencias medicas sobre plataforma web

En esta aplicación se recogen elementos fundamentales que permiten simular, en una mayor escala, el proceso de enseñanza y aprendizaje en un entorno web, donde se integran la programación del lado del

cliente (Javascript) y del servidor (PHP). El Sistema Gestor de Base de Datos que utiliza es MYSQL. (Rodríguez, 2007)

La aplicación permite tener concentrada toda la información referente a los casos que tienen un determinado diagnóstico, además tener unidas todas las características de los rasgos sobre un juicio estipulado; de forma tal que los usuarios pueden consultar todos los rasgos que influyen en la determinación de un diagnóstico y complementar la información de dichos rasgos con el apoyo de imágenes fijas (ejemplos: radiografías, gammagrafías, etc.), imágenes animadas (ejemplos: ecografías, video-EEG, angiografías, etc.) y sonidos (ruidos y soplos cardiacos, respiratorios o de otro tipo). (Rodríguez, 2007)

Esta aplicación está destinada para estudiantes de medicina, enfermería, estomatología, residentes y médicos. Con frecuencia, si se dispone solo de las fuentes tradicionales para hacer alguna consulta acerca de un determinado diagnóstico, se requiere revisar varias fuentes bibliográficas: textos, reproducciones gráficas y otros tipos de información que no es posible en ocasiones recopilar. El sistema incluye un módulo de ejercicios donde se presentan un conjunto de preguntas de tipo reproductivas (Verdadero o Falso, Selección Única, Selección Múltiple, Completar, Enlazar, etc.), y aplicativas; confeccionadas de manera que se puede comprobar la mayor parte del conocimiento mostrado en el sistema. Esto ayuda al estudiante a comprobar todo lo estudiado y también, a que el profesor monitoree el comportamiento evolutivo del alumno. (Rodríguez, 2007)

### **1.1.3 Sistema experto para el diagnóstico neurológico (NDS)**

Sistema experto para el Diagnóstico Neurológico (Neurology Diagnosis System NDS por sus siglas en inglés), es un sistema experto basado en web para el diagnóstico de trastornos neurológicos o los trastornos del sistema nervioso. Este permite a los asistentes de salud, que radican en zonas distantes poder utilizar el sistema para diagnosticar a los pacientes neurológicos, en ausencia de los neurólogos. En cuanto a la metodología, el principio básico es para codificar el conocimiento del dominio de neurología en forma de reglas y casos representativos; utilizando este conocimiento para resolver nuevos casos. Esta codificación del conocimiento es esencialmente la aplicación de dos técnicas de razonamiento llamado, razonamiento basado en casos y basado en reglas. (NDS, 2009)

El resultado de realizar razonamiento basado en reglas es una lista de probables enfermedades y la salida de razonamiento basado en casos, es una lista de casos similares al que será diagnosticado. El componente de razonamiento basado en casos, representa una máquina de aprendizaje que aprende a medida que se introducen nuevos casos. Este sistema experto híbrido, fue construido utilizando el lenguaje de programación Java con MySQL como sistema de base de datos. (NDS, 2009)

## **1.1.4 Sistema Inteligente para la Selección de Información**

El Sistema Inteligente para la Selección de Información (SISI), es un sistema en el cual se logra implementar el Razonamiento Basado en Casos orientado a tareas de diagnóstico. Los casos en SISI se toman de artículos de ficheros de base de datos relacionales. Para realizar el razonamiento se utiliza una red neuronal, la cual se construye a partir del conocimiento de la Base de Casos y se emplea para agilizar este proceso. Los pesos calculados por el algoritmo de aprendizaje de la red, se utilizan en la función de semejanza para la recuperación de los casos más parecidos. Toda esta estructura se programó utilizando el paradigma orientado a objetos. Tiene como objetivo formular una conclusión diagnóstica en psiquiatría a un problema clínico complejo y lograr la eficiencia del experto para el diagnóstico, teniendo como referencia el emitido por los especialistas. (Lio, 1998)

A pesar de que muchos de estos Sistemas Expertos se encuentran en funcionamiento y tienen un buen desempeño partiendo de sus potencialidades, solo responden a las necesidades particulares de las instituciones para los cuales fueron creados. Están desarrollados bajo tecnologías privativas, lo que impide acceder o modificar su código. Y la máquina de inferencia sobre la cual fueron creados no se puede ajustar a otro tipo de información.

## **1.2 Sistemas Expertos**

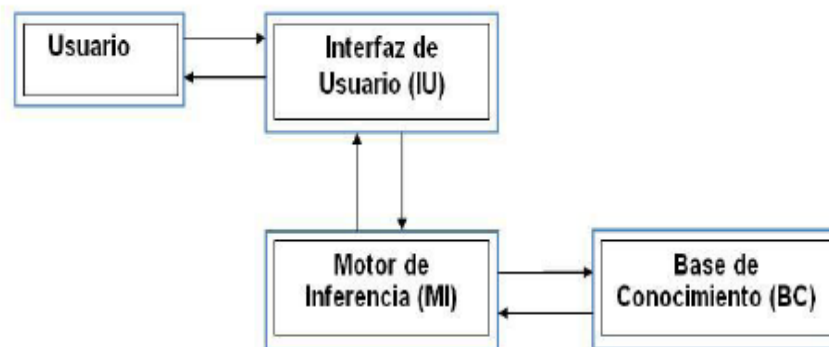
Desde el surgimiento de los Sistemas Expertos (SE), estos han tenido un gran auge, principalmente en el campo de la salud para la predicción y diagnósticos de pacientes. A continuación se brinda una definición de SE:

"Un programa inteligente de computador, que utiliza conocimiento y procedimientos de inferencia para resolver problemas que son lo suficientemente difíciles y requieren experticia humana significativa para su solución. El conocimiento de un Sistema Experto consiste de hechos y heurísticas". (López, 2010)

## 1.2.1 Estructura de un Sistema Experto

Los SE suelen tener una estructura (Bello, Rafael E. , 2002 pág. 100) la cual se describe a continuación:

- **Base de conocimiento:** contiene los conocimientos relativos a la tarea. Usa formalismos de representación para codificar los conocimientos en la base.
- **Motor de inferencias:** medio por el cual controla y aplican los conocimientos. Permiten que el sistema razone a partir de los datos, noticias o conocimientos de entrada para producir los resultados de salida. Gobiernan el orden en el que el sistema realiza los pasos de razonamiento, acepta entradas y produce salidas.
- **Interfaz de E/S:** para que el usuario proporcione hechos, datos y el sistema responda.



**Figura 1.** Estructura de un Sistema Experto (Bello, Rafael E. , 2002 pág. 100)

Para comprender mejor este tipo de sistemas, se debe conocer que: un caso es la definición completa, clara y precisa de las características de un problema particular, que lo distinguen de otros problemas y las acciones que se deben tomar para su corrección. (Prieto, 2011 p. 24)

## 1.2.2 Clasificación de los Sistemas Expertos más usados en la salud

En la clasificación de los SE más utilizados en la salud se encuentran: los Sistemas Basados en Reglas y los Sistemas Basado en Caso. A continuación se brinda una breve descripción de en qué consisten los mismos.

- **Sistema Basado en Reglas**

Los Sistema Basado en Reglas (SBR) constituyen un tipo de SE que utiliza como forma de representación del conocimiento las reglas de producción. Las reglas utilizan un formato SI – ENTONCES (IF – THEN) para representar el conocimiento, la parte SI de una regla es una condición o premisa y la parte ENTONCES de la regla es la acción o conclusión, la cual permite inferir un conjunto de hechos nuevos si se verifican las condiciones de la parte SI. (Bello, Rafael E. , 2002 pág. 100)

El método de solución del problema utilizado por los SBR sigue el método conocido como Módus Ponens, creando una cadena de inferencias, que constituye un camino entre la definición del problema y la solución. Este camino normalmente es creado por una búsqueda primero en profundidad, dirigido por objetivo o datos en dependencia de la cantidad de datos de entrada y de conclusiones. (Bello, Rafael E. , 2002 pág. 100)

- **Sistema Basado en Casos**

Un Sistema Basado en Casos (SBC), es una de las técnicas de Inteligencia Artificial (IA) más empleadas en la actualidad en los procesos de toma de decisiones, son un tipo de SE que utiliza los casos como forma de representación del conocimiento y emplean el Razonamiento Basado en Casos (RBC) como mecanismo de inferencia. El RBC es un proceso que considera casos anteriores para llegar a una decisión, en dependencia de la comparación entre la situación actual y los casos viejos. Dado un nuevo caso para clasificar, un razonador basado en casos verifica primero si existe un caso de entrenamiento idéntico. Si encuentra uno, retorna la solución acompañando a ese caso. Sino encuentra ningún caso idéntico, el razonador buscará casos de entrenamiento que tienen componentes similares a aquellas del nuevo caso. El sistema es capaz de aprender a través de la agregación de nuevos casos. (Bello, Rafael E. , 2002 pág. 100)

### **Justificación a la selección del SBC**

Los SE mencionados anteriormente, presentan como semejanza que son utilizados para la predicción. Pero se definió utilizar el SBC, debido a que existe una base de información respecto a las RAM ocurridas en Cuba, que podría utilizarse como una BC; pudiendo inferir haciendo uso de esta la solución a un problema determinado.

Un SBR será flexible y producirá respuestas casi óptimas, pero sería lento y propenso a errores, debido a que las RAM no se pueden predecir mediante reglas, por la inasistencia de un grupo homogéneo por el cual se podría llegar a una conclusión. El RBC estará restringido a las variaciones sobre situaciones conocidas y producirá respuestas aproximadas, pero será rápido y sus respuestas serán establecidas como experiencia real. Con el SBC siempre hay una conexión corta entre el caso de entrada y la solución recuperada.

- **Estructura de un Sistema Basado en Casos**

En la Base de Casos (BC), se almacena el conocimiento necesario para resolver los problemas del dominio de aplicación, puede ser representada a través de una tabla en la cual las columnas son etiquetadas por variables que representan los rasgos predictores y los rasgos objetivos (decisiones) y las filas representan casos. La Máquina de Inferencia (MI), es un procedimiento basado en un esquema de razonamiento o método de solución de problemas, que utiliza el conocimiento para resolver los problemas de ese dominio y la Interfaz de Usuario (IU) permite la comunicación entre el usuario y el ordenador. (Bello, Rafael E. , 2002 pág. 100)

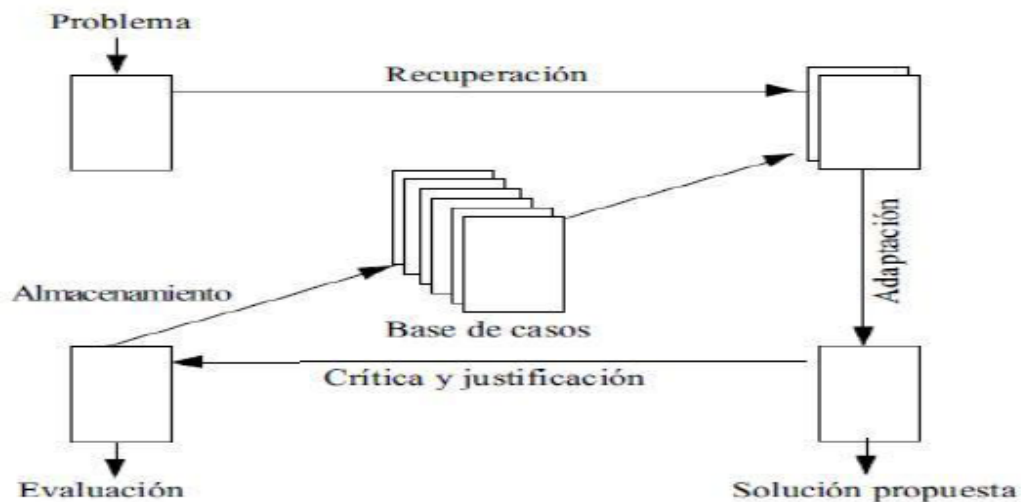
- **Máquina de inferencia en el RBC**

Una parte importante en el RBC, es la etapa de recuperación, la cual se divide en: selección y recuperación. La primera se encarga de seleccionar, entre todos los casos almacenados cuáles son los candidatos que pueden servir para resolver el nuevo problema. La segunda etapa se encarga de recuperar dentro de esos casos seleccionados, cuáles son los más parecidos al nuevo problema, de forma que pueda utilizarse la solución dada a ellos. (Agudo, 2001 pág. 37)

Una vez que se presenta el conjunto de casos candidatos, la recuperación del caso que mayor semejanza tiene con el problema puede producir dos resultados: existe un caso que coincide completamente con el patrón de búsqueda, o varios casos que difieren al menos un rasgo con respecto al patrón de búsqueda. La primera variante significa que el problema estaba almacenado en la Base de Casos y se tiene la solución. En el segundo, existen dos alternativas, una es realizar la búsqueda por analogía y la otra determinar un valor numérico que indique el grado de semejanza entre el problema y cada caso candidato. (Agudo, 2001 p. 37)

## • Ciclo del Razonamiento Basado en Casos

A continuación se muestra la explicación del funcionamiento del ciclo de un RBC, dándose a conocer las definiciones para cada una de las etapas del proceso:



**Figura 2.** Funcionamiento del RBC (Vega, et al., 2000 pág.8)

- La **recuperación** de casos, es la selección en la base de conocimiento de aquellos cuya descripción se ajusta más a la información presentada en el nuevo caso.
- La **adaptación**, consiste en adecuar la solución del caso más parecido a las condiciones del nuevo caso. Esto es necesario, dado que normalmente los fenómenos o síntomas que se presentan en un diagnóstico no son idénticos a los ocurridos en los casos anteriores.
- La etapa de **crítica y justificación**, consiste en la validación de la solución propuesta. Esta validación se realiza contrastando diferentes soluciones o simulando la solución, para estimar qué tan acertada es. Esta etapa está altamente influenciada por el grado de conocimiento que tiene el experto sobre el fenómeno ocurrido y juzga la efectividad de la solución propuesta en base a su experiencia.
- En la etapa de **evaluación**, se aplica la solución propuesta y se analiza el resultado de su aplicación. Si los resultados son los esperados, se confirma la solución, pero si existen diferencias, se debe averiguar por qué ocurrieron las mismas y cómo pueden evitarse. Esta información debe servir para mejorar la definición del caso.



- La etapa de **almacenamiento** consiste en registrar, en la base de conocimiento, la información derivada del nuevo caso, ya sea como un caso nuevo o un caso mejorado.

### 1.2.3 Tipos de búsqueda en los Sistemas Expertos

Durante la recuperación de los casos, se necesita realizar un proceso de búsqueda, el cual brindará como resultados el o los casos que podrán ser tomados como solución al problema. A continuación se expone una breve explicación de los tipos de búsqueda que se pueden utilizar durante el proceso antes mencionado:

- **Búsqueda por analogía**

La búsqueda por analogía, significa determinar si la descripción del problema es equivalente a la descripción de alguno de los casos almacenados, es decir, probar la equivalencia entre los valores de los rasgos predictores, aunque sean diferentes, comprobando si semánticamente tienen igual significado. Para este análisis se utiliza una red semántica, en la cual los nodos representan valores y rasgos. Además los enlaces indican el grado de relación entre ellos. Dos valores son equivalentes si están enlazados con el mismo concepto por el mismo tipo de relación, a partir del mismo punto de vista. (Agudo, 2001 p. 37)

- **Búsqueda por semejanzas**

La búsqueda por semejanzas, tiene como objetivo obtener el caso almacenado que más se parece al nuevo problema. Existen diversas formas de medir el grado de semejanza, estas pueden ser, desde expresiones analíticas de distancia hasta la descripción de algoritmos para obtener el grado de parecido entre dos objetos. Esto depende del tipo de información que se presente. (Rodríguez, 2003 p. 58)

Considerando que un caso está descrito por un número de características, una aproximación sencilla sería contabilizar los atributos comunes entre el caso almacenado y el nuevo problema. Se suele distinguir entre medidas de semejanzas locales y globales, las primeras calculan la semejanza entre valores de un mismo atributo en los dos casos y la segunda combina los resultados de aplicar la similitud local a cada uno de los atributos de los casos que se están comparando. Una estrategia para definir semejanza entre atributos simbólicos, consiste en utilizar una tabla de semejanzas, donde se obtiene explícitamente la similitud entre cada par de valores. (Rodríguez, 2003 p. 58)

Las funciones de semejanza indican cómo se deben comparar los casos, pero es necesario saber qué se debe comparar. De todos los atributos que conforman un caso, es necesario definir cuáles son los más relevantes para efectuar la comparación. Una vez que se decide realizar el proceso de búsqueda utilizando funciones de semejanza, es necesario definir en dependencia del dominio de aplicación, las funciones que pueden emplearse. Existen varias funciones de semejanzas que han sido propuestas para recuperar casos semejantes, muchas de ellas trabajan con atributos numéricos y no son apropiadas para atributos simbólicos y otras permiten buscar semejanza teniendo en cuenta una estructura determinada. (Martínez, et al., 2008 pág. 63)

## **Justificación de la utilización de la búsqueda por semejanza**

Para la recuperación de los casos, se propone utilizar el tipo de búsqueda por semejanza. En el momento de ingresar el caso nuevo se buscará si existe un caso semejante que cumpla con la medida de semejanza local, en caso de no existir, se aplica la semejanza global utilizando la función de semejanza. La propuesta de función de semejanza a utilizar es la realizada por el Dr. Rafael E. Bello Pérez en el trabajo “Un Sistema Basado en Casos para la toma de decisiones en condiciones de incertidumbre”.

### **1.2.4 Incertidumbre en los SBC**

La incertidumbre, siempre está presente en cualquier problema de la vida, debido a que la información incompleta es omnipresente, casi toda la información que se tiene del mundo es incierta, incompleta o imprecisa. El propósito de los sistemas de información es modelar el mundo real y por consiguiente la incertidumbre. (Pérez, y otros, 2002 pág. 107)

La utilización de los SBC, con respecto a otros sistemas, tiene como desventaja que confía ciegamente en los casos previamente almacenados en su memoria. Además puede ser que no recupere el caso más apropiado para la solución del nuevo caso. Estas desventajas disminuyen con la introducción de un mecanismo de manejo de la incertidumbre en estos sistemas. Obviamente el RBC no escapa de la problemática de la incertidumbre. Hay posibilidades de imprecisión, tanto en el conocimiento (Base de Casos) como en el planteamiento del problema. El conocimiento del problema radica esencialmente en la colección de casos y si el caso representa o no el objetivo para el cual fue puesto en la base. (Pérez, y otros, 2002 pág. 107)

## **1.3 Modelos de base de datos**

Se define como base de datos, a una colección o depósito de datos integrados, con redundancia controlada y con una estructura que refleje las interrelaciones y restricciones existentes en el mundo real. Los datos deben mantenerse independientes de las aplicaciones para poder ser compartidos entre ellas. Varias técnicas son usadas para modelar la estructura de datos. La mayor parte de la base de datos son construidas en torno a un modelo de datos en particular. (Avanzadas, 2007 p. 7)

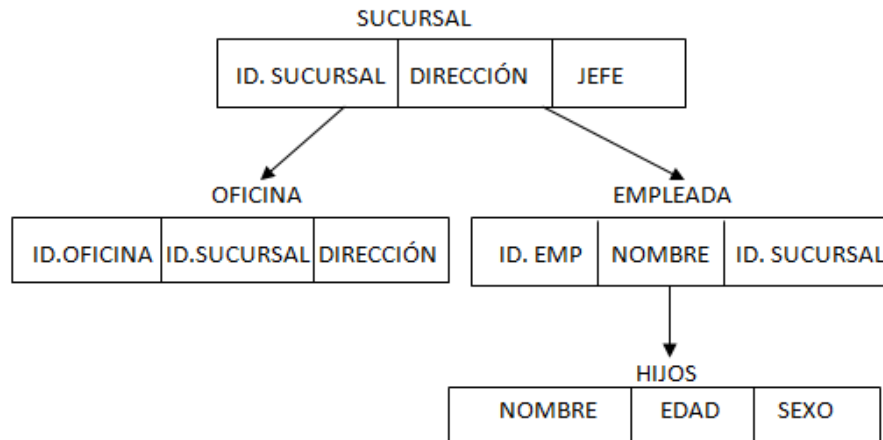
### **1.3.1 Modelo Plana**

El diseño de una base de datos simple o plana, consiste en una tabla grande en lugar de varias tablas interconectadas, llamado "plana" debido a su estructura única de dos dimensiones (campos y registros de datos), estas bases de datos no representan relaciones complejas de datos. Una base de datos plana contiene una o más tablas, cada una conteniendo uno o más campos. Cada tabla es completamente independiente de todas las otras tablas de la base de datos. (Schmieg, 2008)

El modelo de bases de datos planas no es adecuado para la mayoría de las aplicaciones de software en los que hay una necesidad de representar y almacenar relaciones. Sin embargo, algunos desarrolladores de aplicaciones todavía lo utilizan con el fin de reducir el costo y la complejidad de la integración de una base de datos relacional. (Schmieg, 2008)

### **1.3.2 Modelo Jerárquico**

Este modelo utiliza árboles para la representación lógica de los datos. Los árboles están compuestos de unos elementos llamados nodos. El nivel más alto del árbol se denomina raíz. Cada nodo representa un registro con sus correspondientes campos. La representación gráfica de este modelo se realiza mediante la creación de un árbol invertido, los diferentes niveles quedan unidos mediante relaciones. (Lorenzo, 2000 pág. 34)



**Figura 3.** Forma de representación del modelo jerárquico. (Lorenzo, 2000 pág. 34)

En este modelo solo se pueden representar relaciones 1: M (Lorenzo, 2000 pág. 34), por lo que presenta varios inconvenientes:

- No se admiten relaciones N: M.
- Un segmento hijo no puede tener más de un padre.
- No se permiten más de una relación entre dos segmentos.
- Para acceder a cualquier segmento es necesario comenzar por el segmento raíz.
- El árbol se debe de recorrer en el orden designado.

### 1.3.3 Modelo de datos en red

En este modelo las entidades se representan como nodos y sus relaciones son las líneas que los unen. En esta estructura cualquier componente puede relacionarse con cualquier otro. A diferencia del modelo jerárquico, en este modelo, un hijo puede tener varios padres. Los conceptos básicos en el modelo en red (Lorenzo, 2000 pág. 37) son:

- El tipo de registro, que representa un nodo.
- Elemento, que es un campo de datos.
- Agregado de datos, que define un conjunto de datos con nombre.
- Este modelo de datos permite representar relaciones N: M.

## 1.3.4 Modelo de datos relacional

Este modelo es el más utilizado, funciona a partir de tablas bidimensionales para la representación lógica de los datos y sus relaciones.

Algunas de sus principales características (Lorenzo, 2000 pág. 38) son:

- Puede ser entendido y usado por cualquier usuario.
- Permite ampliar el esquema conceptual sin modificar las aplicaciones de gestión.
- Los usuarios no necesitan saber dónde se encuentran los datos físicamente.
- El elemento principal de este modelo es la relación que se representa mediante una tabla.

## Justificación para la selección de un modelo plana

Se decidió para la representación de la información un modelo de base de datos plana, debido a que estas permiten un acceso rápido y sencillo a la información pues esta se encuentra en una sola tabla. La estructura de este tipo de modelo brinda la posibilidad de representar en los registros, cada uno de los casos, permitiendo un acceso rápido a los valores de cada uno de los registros.

## 1.4 Identificación de los rasgos predictores

Para la identificación de los rasgos predictores se propone utilizar el Reconocimiento de Patrones. Este reduce de modo eficiente el número de variables (atributos o rasgos) con los cuales se deben describir los objetos. Además para encontrar los rasgos que inciden de manera determinante en un problema. (Pons, A., 2003 pág. 4)

Una alternativa para la selección de variables, es la utilización del conjunto de Testores Típicos (TT), los cuales se definen como un conjunto de rasgos (columnas) que permite diferencias entre dos clases. Un testor se llama irreducible (típico) si al eliminar cualquiera de dichas columnas deja de ser testor para (T0, T1) (Ochoa, 2004 p. 37). Los TT son determinantes en problemas como: la evaluación de la importancia informacional de los rasgos y la selección de variables, pues permiten reducir la dimensión del espacio de representación de los objetos. (Pons, A., 2003 pág. 4)

### 1.4.1 Algoritmos existentes para el cálculo de los Testores Típicos

Para el cálculo de los testores típicos se han desarrollado varios algoritmos, que por su estrategia de cómputo pueden clasificarse en algoritmos de escala exterior y de escala interior. Los algoritmos de escala exterior, son aquellos que realizan el cálculo de los TT generando los elementos del conjunto

potencia del conjunto de columnas de la Matriz Básica (MB), de tal forma que, usando determinados criterios, el algoritmo trata de evitar el análisis de todos los subconjuntos. Por el contrario, los algoritmos de escala interior realizan el cálculo de los TT basados en el estudio de la estructura interna de la matriz, encontrando condiciones que garantizan la característica de testor y la tipicidad en las columnas asociadas a las posiciones unitarias de la matriz. (Pons, A., 2003 pág. 4)

- **Algoritmo BT**

El algoritmo BT es de escala exterior. El orden que caracteriza a este algoritmo es el generado por los números naturales de forma ascendente en su notación binaria mediante un n-uplo booleano. Este constituye un orden total sobre los elementos no nulos del conjunto potencia del conjunto de rasgos de la Matriz de Aprendizaje (MA). El BT comienza por el n-uplo (0...01) y procede comprobando si cada vector (n-uplo) generado es un testor o un TT. Según el resultado, establece “saltos” en el conjunto potencia. El algoritmo termina cuando llega al n-uplo (1...11). Para ello se apoya en caracterizaciones de los conceptos de testor y TT y en proposiciones, que definen los saltos cuando el n-uplo analizado es o no un testor. (Pons, A., 2003 pág. 4)

- **Algoritmo CT**

El algoritmo CT es de escala interior y trabaja basado en los conceptos de conjunto de filas independientes y conjunto completo. Inicialmente se encuentran los testores y luego, se prueba si son típicos o no. Por definición todo conjunto completo es un testor. El método del CT consiste en explorar las posiciones unitarias de la MB, encontrando los conjuntos completos y comprobando, entonces, si son TT. (Pons, A., 2003 pág. 5)

- **Algoritmo CC**

El algoritmo CC es de escala interior y está basado en los conceptos de Conjunto Compatible y Conjunto Regular. El CC tiene como peculiaridad, a diferencia del CT, que su objetivo es garantizar la tipicidad. Esto está implícito en la definición de conjunto compatible, luego tiene que comprobar si se cumple la propiedad de testor. (Pons, A., 2003 pág. 5)

- **Algoritmos REC y CER**

Ambos son algoritmos de escala exterior. El REC tiene la particularidad de que trabaja directamente con la MA, lo cual lo sitúa en desventaja con relación a los restantes algoritmos por el enorme manejo de

información que tiene que realizar. El CER fue encaminado a resolver este problema y lo hizo introduciendo un nuevo orden en el conjunto potencia de los rasgos. (Pons, A., 2003 pág. 5)

- **Algoritmo Lexicográfico**

El algoritmo Lexicográfico (LEX) va generando listas de rasgos. Construye listas de rasgos que posean la propiedad de tipicidad y luego, comprueba si este conjunto de rasgos constituye un testor. Cuando se obtiene un TT se producen saltos. (Pons, A., 2003 pág. 6)

El algoritmo comienza analizando el primer rasgo de MB. Un nuevo rasgo se puede incorporar a la lista si se cumple que no es excluyente con la lista (puede coexistir con los rasgos de la lista para formar un TT y tiene filas típicas con respecto a la lista). Esta condición es la que garantiza la propiedad de tipicidad. Luego, se comprueba si el conjunto de rasgos contenido en la lista junto con el nuevo rasgo forma un testor, verificando si no existe fila en la MB (considerando sólo esos rasgos) completa de ceros. Si se cumple que es testor, entonces se está en presencia de un TT y se almacena. Si se encontró un TT y éste contiene al último rasgo de MB, entonces se saltan todos los sus subconjuntos consecutivos. Si no contiene al último rasgo de MB, para saltar todos los súper conjuntos del TT encontrado, se elimina el último rasgo de la lista y se analiza si se puede incluir el próximo rasgo de la MB. (ver Anexo 8) (Pons, A., 2003 pág. 6)

## **Justificación a la selección del algoritmo LEX**

Se seleccionó el algoritmo LEX entre los planteados anteriormente para el cálculo del TT, debido que a diferencia de los demás, este disminuye el tiempo necesario para la obtención de todos los TT de una MA. Posee la peculiaridad de que busca los TT implantando un orden sobre el conjunto potencia de los rasgos, definiendo ciertos saltos con el objetivo de obviar algunos conjuntos.

## **1.5 Herramientas y tecnologías a utilizar**

Para desarrollar el SBC, se utilizarán las herramientas y tecnologías propuestas por el proyecto Sistema para el control Farmacológico (Synta), pues la solución se encontrará dentro de dicho sistema.

### **1.5.1 Marcos de desarrollo**

Los marcos de desarrollo (*frameworks*) se pueden considerar como soluciones completas que contemplan herramientas de apoyo a la construcción (ambiente de trabajo o desarrollo) y motores de ejecución

(ambiente de ejecución). Estos aceleran el proceso de desarrollo, permiten reutilizar código ya existente y promover buenas prácticas de desarrollo, como el uso de patrones. (Zaninotto, y otros, 2008 pág. 11)

## 1.5.2 Symfony 1.4

Es un *framework* de desarrollo web, que simplifica la creación de una aplicación, mediante la automatización de algunos de los patrones utilizados para resolver tareas comunes. Proporciona estructura al código fuente, forzando al desarrollador a crear códigos más legibles y fáciles de mantener. Posibilita la programación de sistemas y encapsula operaciones complejas en instrucciones sencillas. Symfony separa la lógica de negocio, la lógica de servidor y la presentación del sistema. Proporciona varias herramientas y clases, encaminadas a reducir el tiempo de desarrollo de una aplicación web compleja. Además, automatiza las tareas más comunes, permitiendo al desarrollador dedicarse por completo a los aspectos específicos de cada aplicación. (Zaninotto, y otros, 2008 pág. 11)

Características de Symfony (Zaninotto, y otros, 2008 pág. 11):

- Fácil de instalar y configurar en la mayoría de las plataformas
- Independiente del sistema gestor de base de datos
- Código fácil de leer, permitiendo un mantenimiento sencillo
- Permite la integración con librerías desarrolladas por terceros

## 1.5.3 HTML

HyperText Mark-Up Language (HTML), se conoce como "lenguaje de marcado", cuya función es preparar documentos escritos aplicando etiquetas de formato. HTML se usa también para la lectura de documentos en Internet desde diferentes equipos, gracias al protocolo HTTP, que permite a los usuarios acceder de forma remota a documentos almacenados en una dirección específica de la red. (HTML, 2013)

## 1.5.4 CSS

Cascade Style Sheet (CSS), traducido significa: Hojas de Estilo en Cascada. Permiten controlar la apariencia de una página web; describiendo cómo los elementos dispuestos en la página son presentados al usuario. CSS es un gran avance que complementa el HTML y es un estándar aceptado por toda la industria relacionada con la Web, o por lo menos, gran parte de los navegadores. (CCS, 2012)



## 1.5.5 JavaScript 1.1

JavaScript es un lenguaje script u orientado a documento, como pueden ser los lenguajes de macros, que tienen muchos procesadores de texto y planillas de cálculo. JavaScript es una de las múltiples maneras que han surgido para extender las capacidades del lenguaje HTML, al ser la más sencilla, es por el momento la más extendida. No se puede desarrollar un programa con JavaScript que se ejecute fuera de un navegador, es un lenguaje interpretado que se embebe en una página web HTML. (Javascript, 2009)

## 1.5.6 ExtJS 3.1

ExtJS es una librería JavaScript que mantiene el código bien estructurado. Incluso las aplicaciones más grandes pueden ser de fácil mantenimiento. Ofrece una recopilación de elementos de interfaz de usuario con temas elegantes de partida. Una potencia de la base de ExtJS, es su diseño de componentes, que permite ampliar fácilmente los componentes predeterminados para satisfacer las necesidades del cliente. Utiliza tecnologías de navegación específica, para proporcionar una experiencia consistente y de alto rendimiento al usuario. Hace un amplio uso de imágenes para la tematización, y de tecnologías específicas para el manejo de datos. (Sencha, 2011)

## 1.5.7 Lenguaje de Programación

HyperText Pre-Processor (PHP, por sus siglas en inglés), es un lenguaje de programación interpretado, diseñado originalmente para la creación de páginas web dinámicas. (PHP, 2012)

Características (PHP, 2012):

- Es un lenguaje multiplataforma
- Completamente orientado al desarrollo de aplicaciones web dinámicas con acceso a información almacenada en una base de datos
- El código fuente escrito en PHP es invisible al navegador y al cliente. El servidor es el que se encarga de ejecutar el código y enviar su resultado HTML al navegador. Esto hace que la programación en PHP sea segura y confiable.
- Capacidad de conexión con la mayoría de los motores de base de datos que se utilizan en la actualidad. Destaca su conectividad con MySQL y PostgreSQL.
- Es libre, por lo que se presenta como una alternativa de fácil acceso para todos
- Permite aplicar técnicas de programación orientada a objetos

## 1.5.8 NetBeans 6.9

NetBeans IDE es una aplicación de código abierto. Diseñada para el desarrollo de aplicaciones fácilmente portables entre las distintas plataformas. Hace uso de la tecnología Java y en ella se pueden desarrollar aplicaciones web utilizando varios lenguajes de programación, entre ellos PHP. Brinda la posibilidad de incorporar el *framework* Symfony, que permite la optimización del desarrollo de sistemas web. (Netbeans, 2010)

## 1.5.9 Servidor Web Apache 2.2

Representa el complemento perfecto para el desarrollo de páginas dinámicas con PHP. Caracterizado principalmente por ser gratuito, de alta popularidad, sencillez de manejo y versatilidad. Puede ser instalado en plataformas Linux y Windows. (Ciberaula, 2010)

Algunas de las características (Ciberaula, 2010) que presenta son:

- *Fiabilidad*: un gran número de los servidores con más alta disponibilidad funcionan bajo un servidor Apache.
- *Software libre*: el servidor Apache es totalmente gratuito y es distribuido bajo la licencia de Apache, en la cual se permite realizar cambios al código fuente.
- *Extensibilidad*: se pueden añadir módulos para ampliar las capacidades del servidor.

## 1.5.10 Sistema gestor de base de datos MySQL 5.0

El software MySQL proporciona un servidor de base de datos SQL (Structured Query Language) muy rápido, multi-threaded, multi usuario y robusto. El servidor MySQL está diseñado para entornos de producción críticos, con alta carga de trabajo. Entre sus principales características se encuentran que es un servidor de base de datos muy rápido, fiable y fácil de usar. (MySql, 2011)

## 1.5.11 EMS MySQL Manager 3.7

Es una potente herramienta para la administración y mantenimiento de un servidor MySQL. Posee una interfaz gráfica de usuario que posibilita creación/edición de base de datos MySQL de una manera simple y sencilla. Este soporta Script SQL, construcciones visuales de consulta SQL, extrae o imprime metadatos y exporta/importa datos. Estos servicios permiten que el trabajo con MySQL sea más práctico, agradable y sencillo. (SQL Manage, 2013)

## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

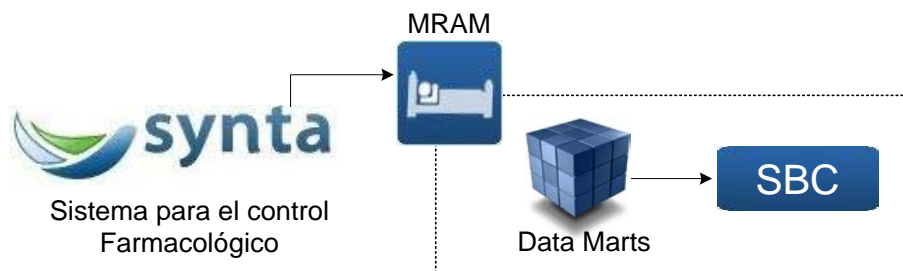
El objetivo fundamental de este capítulo es abordar los aspectos tenidos en cuenta en la realización del Sistema Basado en Casos, describiendo el flujo seguido durante la implementación.

### 2.1 Propuesta de solución

La Universidad de las Ciencias Informáticas (UCI) que está constituida por centros productivos. Uno de estos es el Centro de Informática Médica (CESIM), encargado de desarrollar aplicaciones para el sector de la salud. El Departamento Sistemas de Apoyo a la Salud (SAS) del CESIM, desarrolla un producto llamado Synta: Sistema para el Control Farmacológico. Este sistema está compuesto por un Módulo de Reacciones Adversas a Medicamentos (MRAM), que permite el registro de la información referente a las reacciones adversas ocurridas en Cuba, así como crear gráficas de comparación entre medicamentos y realizar consolidados de información. A pesar que en el módulo se almacena la información de las RAM, los médicos no pueden hacer uso de esta para realizar posibles predicciones (sobre la ocurrencia de las RAM) teniendo en cuenta varios indicadores.

Por lo que se propone, implementar un SBC e integrarlo al MRAM, agregándole una nueva funcionalidad al módulo. Para el llenado de la Base de Casos, se propone hacer uso del Data Mart que tiene dicho módulo. Esta integración le permitirá al SBC tener registrados casos únicos (estos casos no presentarán información incompleta), logrando agilizar la búsqueda de información.

A continuación se muestra una figura representando la propuesta del flujo de información.

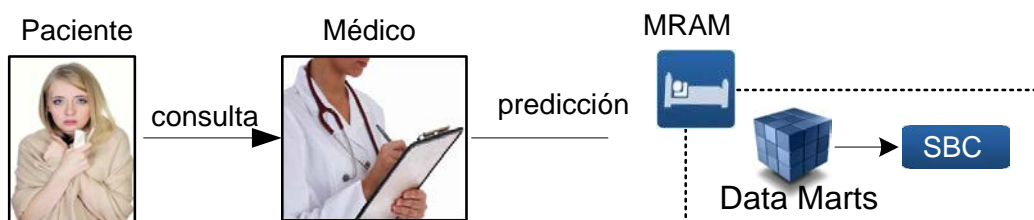


**Figura 4.** Integración del Sistema Basado en Casos en el MRAM del producto Synta (elaboración propia)

# Capítulo 2: Propuesta de solución del Sistema Basado en Casos

El paciente acude a la consulta médica, el médico realiza la prescripción de los medicamentos teniendo en cuenta los padecimientos que este presenta. Para predecir que RAM le podría causar al paciente por consumir dichos medicamentos, el médico debe introducir los datos necesarios respecto a este paciente en el SBC. El sistema analiza los datos introducidos y realiza la predicción, mostrando el nivel de certeza asociado al mismo e indicando las posibles RAM que le podría causar.

El flujo a seguir para la utilización del SBC durante la consulta médica quedaría de la siguiente forma:



**Figura 5.** Propuesta de flujo a seguir para predecir una RAM desde el MRAM (elaboración propia)

## 2.2 Técnicas utilizadas en el desarrollo del SBC

Para la creación del SBC se aplicaron distintas técnicas, estas se explican a continuación en el orden en que fueron utilizadas y se muestra en qué componentes incidieron. (ver Figura 6)

	Pasos	Componente del SBC
1.	Determinación de los Testores Típicos	} Base de Casos
2.	Cálculo del peso de los rasgos predictores mediante el Triángulo de Fuller	
3.	Determinación de la Incertidumbre	} Máquina de Inferencia
4.	Calculo de la función de semejanza	

**Figura 6.** Pasos para la implementación del SBC (elaboración propia)

### Paso 1. Determinación de los Testores Típicos

Para el cálculo de los Testores Típicos, se utilizó el algoritmo LEX, el cual a partir de setenta y cuatro atributos (rasgos) relacionados a la información almacenada referente a las RAM, permitió obtener un

## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

conjunto de Testores Típicos. Estos rasgos se dividieron en 20 objetos y 3 clases, obteniéndose una Matriz Básica (ver Figura 7) que posibilitó obtener los testores típicos.

11110000  
10101000  
10010010  
10001011  
10000100  
11001010

**Figura 7.** Matriz Básica

Los testores típicos obtenidos fueron:

- (grupo etario)
- (fabricante,sexo,raza,municipio,app)
- (fabricante,sexo,Medicamento sosp,motivo prescr)
- (fabricante,raza,motivo prescr)
- (fabricante,municipio,Medicamento sosp)

Los cuales fueron mostrados al experto, determinándose así seis rasgos, que constituirían los rasgos predictores a utilizar en la Máquina de Inferencia del SBC; los cuales se describen a continuación:

- *Motivo de prescripción:* motivo por el cual el médico realiza la prescripción
- *Grupo etario:* las edades son agrupadas por rango de edades, formando tres grupos de la siguiente manera: 0-15 años niño (n), 16-65 años adultos (a) y (65 y +); geriátrico (g)
- *Medicamento sospechoso:* medicamento a prescribir al paciente
- *Antecedente patológico personales (APP):* antecedentes médico-quirúrgicos, traumatismos y transfusiones sanguíneas que presenta el paciente
- *Fabricante:* fabricante del medicamento sospechoso
- *Sexo:* sexo del paciente
- *Municipio:* lugar donde reside el paciente

Estos son los atributos que conformarán las tablas de la BC del SBC.

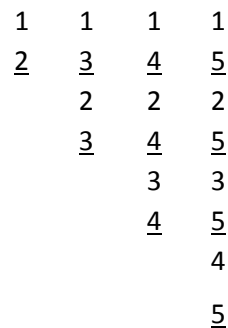
# Capítulo 2: Propuesta de solución del Sistema Basado en Casos

## Paso 2. Cálculo del peso de los rasgos predictores mediante el Triángulo de Fuller

El Método del Triángulo de Fuller, se utilizó para obtener la importancia o peso de los rasgos predictores (Rodríguez, 2010 pág. 209). Para calcular el peso de los rasgos, primeramente fue necesario enumerarlos (Quintana, y otros, 2004), quedando de la siguiente manera:

- 1- Medicamento sospechoso
- 2- Motivo de prescripción
- 3- Fabricante
- 4- APP
- 5- Municipio

Los rasgos sexo y grupo etario son rasgos compuestos, por lo que se le aplicó la técnica del Triángulo de Fuller de forma independientemente. Una vez enumerados los rasgos se construye el triángulo de Fuller, utilizando la numeración realizada anteriormente, como se muestra a continuación:



**Figura 8.** Creación del triángulo de Fuller

Seguidamente se hace una comparación (en dúo) entre cada uno de los rasgos, marcando con un asterisco rojo el que tiene mayor importancia o relevancia según el experto, como se muestra en la siguiente figura:

1*	1*	1	1*
2	3	4*	5
	2*	2	2*
	3	4*	5
		3	3*
		4*	5
			4*
			5

**Figura 9.** Señalamiento de los rasgos más importantes según criterio del experto

Una vez los rasgos más importantes, se calcula el valor de  $\lambda$  para cada uno de ellos, representando la cantidad de ocasiones que cada número es marcado con asterisco rojo.

Para el rasgos 1,  $\lambda_1 = 3$

Para el rasgos 2,  $\lambda_2 = 2$

Para el rasgos 3,  $\lambda_3 = 1$

Para el rasgos 4,  $\lambda_4 = 4$

Para el rasgos 5,  $\lambda_5 = 0$

Luego se procede al cálculo del valor de cada rasgo predictor aplicando la fórmula:

$$V_t = \lambda t / N \text{ (Quintana, y otros, 2004 pág. 9)}$$

Donde  $N$  se obtiene a partir de:

$$N = k(k - 1) / 2 \text{ (Quintana, y otros, 2004 pág. 9)}$$

Donde  $k$  representa la cantidad total de rasgos predictores, sustituyendo los valores quedaría:

$$N = \frac{5(5 - 1)}{2} = 10$$

Al calcular el valor de cada uno de los rasgos, se obtuvieron los siguientes resultados:

$$V_1 = \frac{3}{10} = 0.3 \quad V_2 = \frac{2}{10} = 0.2 \quad V_3 = \frac{1}{10} = 0.1 \quad V_4 = \frac{4}{10} = 0.4 \quad V_5 = \frac{0}{10} = 0$$

## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

En los casos de sexo y grupo etario se obtuvieron los siguientes pesos:

Sexo: F= 0,9 M= 0,1

Grupo: niño = 0,6 adulto = 0,1 geriátrico = 0,3

Obtenidos estos resultados quedaron definidos los pesos para cada uno de los rasgos predictores, en el caso del Municipio al tener peso cero, se toma como un dato opcional a tener presente a la hora de mostrar el resultado.

### Paso 3. Determinación de la certidumbre

#### Formación de las tablas de contingencia, determinación de Chi-cuadrado y V-Cramer

La Base de Casos, se segmenta en tablas de contingencias para determinar el estadístico Chi-cuadrado, el cual es utilizado como probabilidad para comparar dos o más grupos de sujetos con respecto a una variable categórica. En otros términos este se utiliza para medir qué grado de asociación que tienen los rasgos predictivos con respecto a los rasgos objetivos. La tabla de contingencia es un medio particular de representar simultáneamente dos caracteres observados en una misma población, si son discretos o continuos o reagrupados en clases. (Pérez, y otros, 2002 pág. 110)

Como complemento al estadístico Chi-cuadrado se utiliza el estadístico V-Cramer para determinar el grado de asociación de un rasgo predictivo ante el rasgo objetivo. Es necesario utilizar dicho complemento, debido a que el resultado que ofrece Chi-cuadrado en presencia de un rasgo objetivo que presente más de dos respuestas asociadas, no se comporta de la misma manera que con más valores. La frecuencia esperada es un valor necesario para calcular Chi-cuadrado. A continuación se ilustra cómo se calcula la misma (Pérez, y otros, 2002 pág. 110):

$$E_{pq}(x_i) = \frac{\sum_{k=1}^n O_{kq}(x_i) * \sum_{j=1}^{\theta} O_{pj}(x_i)}{n}$$

**Fórmula 1:** Fórmula para obtener las frecuencias esperadas (Pérez, y otros, 2002 pág. 110)



## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

- **Cálculo de Chi-Cuadrado**

Las diferencias  $O_{pq}(x_i) - E_{pq}(x_i)$ , se llaman residuales, se elevan al cuadrado para evitar la compensación de diferencias positivas negativas. Estas se dividen por la frecuencia esperada  $E_{pq}(x_i)$ , para establecer magnitudes relativas. La fórmula que representa el estadístico es:

$$x^2(x_i) = \sum_{p=1}^n \sum_{q=1}^{\theta} \frac{(O_{pq}(x_i) - E_{pq}(x_i))^2}{E_{pq}(x_i)}$$

**Fórmula 2:** Fórmula para obtener el valor de Chi-Cuadrado (Pérez, y otros, 2002 pág. 110)

- **Cálculo de V-Cramer**

El estadístico V-Cramer se utiliza para determinar el grado de asociación.

$$V(x_i) = \sqrt{\frac{x^2(x_i)}{n * (k - 1)}}$$

**Fórmula 3:** Fórmula para obtener el valor de V-Cramer (Pérez, y otros, 2002 pág. 111)

Donde  $k = \min$  (número de filas, número de columnas)

Cuanto mayor sea el valor de V, mayor será el grado de asociación.

Si el valor  $V(x_i) \geq \alpha$  (nivel prefijado de antemano) la asociación es considerada. El mejor predictor es aquel que tenga el mayor valor de  $V(x_i)$ . (Pérez, y otros, 2002 pág. 111)

### Árboles de decisión

La utilización de árboles de decisión posee algunos inconvenientes o limitaciones, una de ellas es la altura del árbol, la cual está condicionada por el número de rasgos predictores que posea la Base de Casos. Para reducir este inconveniente, el árbol es minimizado cuando se aplica la técnica de separación de rasgos basada en el estadístico V-Cramer y a un umbral definido como  $\alpha$ , que indica que si el estadístico es mayor que el valor de  $\alpha$  entonces el rasgo se considera asociado. En cualquier otro caso este se considera no asociado, siendo los asociados los únicos que formarán el árbol de decisión.

# Capítulo 2: Propuesta de solución del Sistema Basado en Casos

## Determinación de las probabilidades

Una vez representada la Base de Casos a través de un árbol de decisión, se procede a calcular la certidumbre asociada a los valores de los rasgos predictivos y objetivos. Esta certidumbre se calcula a partir de dos magnitudes conocidas como probabilidad 1 y probabilidad 2 (P1 y P2).

- **Determinación de la primera probabilidad**

La P1 se calcula usando el árbol de decisión que describe, la interacción entre los valores de los rasgos predictivos y objetivos. Para el caso de los rasgos predictivos que se consideran no asociados, la probabilidad P1 se calcula como su frecuencia de aparición en la Base de Casos. A continuación se ilustra la fórmula para calcular la primera magnitud presente en los rasgos no asociados, asociados y rasgos objetivos:

$$P_1(x_i(O_t) = x_{ip}) = \frac{|x_{ip}|}{n}$$

**Fórmula 4:** Fórmula para obtener la primera probabilidad de los rasgos predictivos no asociados (Pérez, y otros, 2002 pág. 113)

$$P_1(x_i(O_t) = x_{ip}) = \frac{|l_{ib}^q|}{\left| \begin{matrix} l^q \\ l_{-1} \end{matrix} \left[ \frac{b}{\eta_i} \right] \right|}$$

**Fórmula 5:** Fórmula para obtener la primera probabilidad de los rasgos predictivos asociados (Pérez, y otros, 2002 pág. 113)

$$P_1(y_1(O_t) = y_{1q}) = \frac{|l_{ib}^q|}{\left| \bigcup_{j=1}^{\theta} l_{ib}^j \right|} \prod_{x_i \in K^0} P(x_i(O_t) = x_{ij})$$

**Fórmula 6:** Fórmula para obtener la primera probabilidad de los rasgos objetivos (Pérez, y otros, 2002 pág. 114)

- **Determinación de la segunda probabilidad**

La P2 representa la certidumbre del valor conocida (certeza total, criterios de expertos e instrumentos de medición). La misma puede variar de uno a otro rasgo predictivo. A continuación se muestra cómo se determina la segunda probabilidad en caso de ser tomada por una fuente con certeza total, instrumentos de medición y por expertos:

$$P_2(x_i(O_t) = x_{ij}) = 1$$

**Fórmula 7:** Fórmula para obtener la segunda probabilidad siendo este valor tomado de una fuente confiable (Pérez, y otros, 2002 pág. 114)

$$P_2(x_i(O_t) = x_{ij}) = \frac{\sum_{r=1}^n c_r * a_r(x_{ij})}{\sum_{r=1}^n c_r}$$

**Fórmula 8:** Fórmula para obtener la segunda probabilidad siendo este valor tomado a partir del criterio de los expertos involucrados con el sistema (Pérez, y otros, 2002 pág. 114)

$$P_2(x_i(O_t)x_{ij}) = P$$

**Fórmula 9:** Fórmula para obtener la segunda probabilidad siendo este valor tomado de un instrumento de medición. (Pérez, y otros, 2002 pág. 114)

$$P_2(y_1(O_t) = y_{1k}) = \frac{\sum_{i=1}^n P_2(x_i(O_t) = x_{ij})}{n}$$

**Fórmula 10:** Fórmula para calcular la segunda probabilidad para el rasgo objetivo (Pérez, y otros, 2002 pág. 114)

- **Determinación de la certidumbre total**

Por último se calculó la certidumbre total de cada valor del rasgo predictor y del rasgo objetivo para cada caso, usando una de las funciones de combinación que han sido utilizadas en la literatura para propagar incertidumbre y que se conocen como co-t normas arquimedeanas por ejemplo:

$$g(x, y) = x + y - x * y$$

**Fórmula 11:** Fórmula para obtener la certidumbre total de cada caso (Pérez, y otros, 2002 pág. 114)

#### **Paso 4. Cálculo de la función de semejanza**

Una vez calculada la certidumbre de todos los valores para los rasgos predictivos, objetivos y la certidumbre de los valores del nuevo caso, se procede a realizar los procesos de razonamiento (recuperación y adaptación). El procedimiento de recuperación consta de dos etapas: acceso y recuperación. Para el sistema propuesto se utiliza la siguiente función de semejanza:

# Capítulo 2: Propuesta de solución del Sistema Basado en Casos

$$\beta(O_0O_t) = \frac{\sum_{i=1}^n p_i * \delta(O_0O_t)}{\sum_{i=1}^n p_i}$$

**Fórmula 12:** Función de semejanza utilizada en la máquina de inferencia (Pérez, y otros, 2002 pág. 115)

La función de similitud se muestra a continuación:

$$\delta'_i(O_0O_t) = \begin{cases} \delta_i * \frac{\xi_i(O_0) + \xi_i(O_t)}{2} & \text{si } (\xi_i(O_0) \wedge \xi_i(O_t)) \geq \varepsilon \\ 0 & \text{e. o. c} \end{cases}$$

**Fórmula 13:** Función de similitud. (Pérez, y otros, 2002 pág. 115)

## Determinación de la decisión

Para este sistema, la adaptación se efectúa determinando el caso de mayor utilidad esperada. Esto se realiza a través de la fórmula siguiente:

$$\mu(O_t) = \gamma\beta(O_0, O_t) + (1 - \gamma) * v_1(O_t)$$

**Fórmula 14:** Fórmula para determinar la certidumbre total (Pérez, y otros, 2002 pág. 115)

## 2.3 Flujo del modelo a implementar

El sistema propuesto consta de seis rasgos predictores y un rasgo objetivo, los rasgos predictores son: grupo etario, sexo, medicamento sospechoso, fabricante, motivo de prescripción y antecedente patológico. En la tabla mostrada a continuación se le asignan a los rasgos predictores un acrónimo:

**Tabla 1:** Rasgos Predictivos (elaboración propia)

Acrónimo	Rasgo Predictor
X <sub>1</sub>	Grupo Etario
X <sub>2</sub>	Sexo
X <sub>3</sub>	Medicamento Sospechoso
X <sub>4</sub>	Fabricante
X <sub>5</sub>	Motivo Prescripción
X <sub>6</sub>	Antecedente Patológico

A continuación se muestra un fragmento de la Base de Casos con datos que serán usados, para representar lo más explícito posible el flujo de implementación seguido en la propuesta de sistema:

# Capítulo 2: Propuesta de solución del Sistema Basado en Casos

**Tabla 2:** Tabla de que representa la Base de Casos (elaboración propia)

Casos	Grupo Etario	Sexo	Fabricante	Ram Principal
	$X_1$	$X_2$	$X_3$	$Y_1$
$O_1$	g	F	MedSol	Cefalea
$O_2$	n	F	MedSol	Vértigo
$O_3$	g	M	Novatec	Vértigo
$O_4$	g	M	MedSol	Cefalea
$O_5$	n	F	Novatec	Vértigo
$O_6$	n	F	Novatec	Vértigo
$O_7$	n	F	MedSol	Vértigo
$O_8$	g	F	MedSol	Cefalea
$O_9$	g	M	Novatec	Cefalea
$O_{10}$	g	M	MedSol	Vértigo

Pasos seguidos durante el flujo de implementación:

**Primer paso:** Creación de las tablas de contingencia

Se deben crear tantas tablas de contingencia como rasgos predictivos existan, luego estos valores serán utilizados en el cálculo de los estadísticos Chi-Cuadrado y V-Cramer, mediante las fórmulas 1, 2 y 3 respectivamente.

**Tabla 3:** Tabla de contingencia para la relación  $X_1/Y_1$  (elaboración propia)

$X_1/Y_1$	Cefalea	Vértigo	Total
g	4	2	6
n	0	4	4
	4	6	10

**Tabla 4:** Tabla de contingencia para la relación  $X_2/Y_1$  (elaboración propia)

$X_2/Y_1$	Cefalea	Vértigo	Total
F	2	4	6
M	2	2	4
	4	6	10

**Tabla 5:** Tabla de contingencia para la relación  $X_3/Y_1$  (elaboración propia)

## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

$X_3/Y_1$	Cefalea	Vértigo	
MedSol	3	3	6
Novatec	1	3	4
	4	6	10

Las tablas anteriores muestran la cantidad de veces que un rasgo predictor incidió sobre un rasgo objetivo. Para el caso del grupo etario X1, con el valor g, existieron cuatro casos que desenlazaron en Cefalea como reacción adversa. Se debe analizar estas relaciones para el resto de los valores y cada uno de los desenlace.

Luego de la confección de las tablas de contingencia se procede al cálculo de los estadísticos para cada uno de los rasgos, teniendo en cuenta el total de filas y columnas de las tablas, a continuación se muestra un ejemplo:

**Tabla 6:** Ejemplo de tabla de contingencia (elaboración propia)

$X_1/Y_1$	Cefalea	Vértigo	Total
g	4 (a)	2 (b)	6 (a+b)
n	0 (c)	4(d)	4 (c+d)
	4 (a+c)	6 (b+d)	10 (n)

**Segundo paso:** se procede al cálculo de las frecuencias esperadas mediante el uso de las fórmulas que se muestran a continuación:

$$E_{11} = \frac{(a + b) * (a + c)}{n} \quad E_{21} = \frac{(c + d) * (a + c)}{n}$$

$$E_{12} = \frac{(a + b) * (b + d)}{n} \quad E_{22} = \frac{(c + d) * (b + d)}{n}$$

**Fórmula 15:** Fórmulas para el cálculo de las frecuencias esperadas (Pérez, y otros, 2002 pág. 110)

**Tercer Paso:** Una vez obtenidas las frecuencias esperadas se procede al cálculo del estadístico Chi-Cuadrado:

$$x^2 = \frac{(a - E_{11})^2}{E_{11}} + \frac{(b - E_{12})^2}{E_{12}} + \frac{(c - E_{21})^2}{E_{21}} + \frac{(d - E_{22})^2}{E_{22}}$$

**Fórmula 16:** Fórmula para el cálculo del estadístico Chi-Cuadrado. (Pérez, y otros, 2002 pág. 110)

## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

Obteniéndose los resultados que se muestran en la siguiente tabla:

**Tabla 7:** Valores obtenidos en cada uno de los rasgos (elaboración propia)

Rasgo	Chi-Cuadrado
X <sub>1</sub>	4,43
X <sub>2</sub>	0,26
X <sub>3</sub>	0,63

**Cuarto paso:** Una vez obtenidos los valores para Chi-Cuadrado, estos serán utilizados para el cálculo del estadísticos V-Cramer para cada uno de los rasgos predictivos mediante la siguiente fórmula:

$$V(x_i) = \sqrt{\frac{x^2(x_i)}{n * (k - 1)}}$$

**Fórmula 17:** Fórmula para la determinación del estadístico V-Cramer (Pérez, y otros, 2002 pág. 111)

Al ser aplicada a cada uno de los rasgos se obtuvieron los siguientes resultados:

**Tabla 8:** Valores obtenidos para cada uno de los rasgos (elaboración propia)

Rasgo	V-Cramer
X <sub>1</sub>	0,66
X <sub>2</sub>	0,16
X <sub>3</sub>	0,25

**Quinto paso:** Definición de los rasgos en asociados y no asociados

Para la definición de los rasgos en asociados y no asociados, se tuvo en cuenta el valor  $\alpha$ . Este valor definió cuan alto sería el nivel de exigencia para la asociatividad entre los rasgos predictores y objetivos. Dado que el valor de  $\alpha = 0.1$  y teniendo en cuenta los valores de V-Cramer obtenidos (grupo etario con 0.66, sexo con 0.16 y fabricante con 0.25) para los rasgos predictores, se definió que: grupo etario, sexo y fabricante son rasgos asociados.

**Sexto paso:** Creación de los árboles de decisión

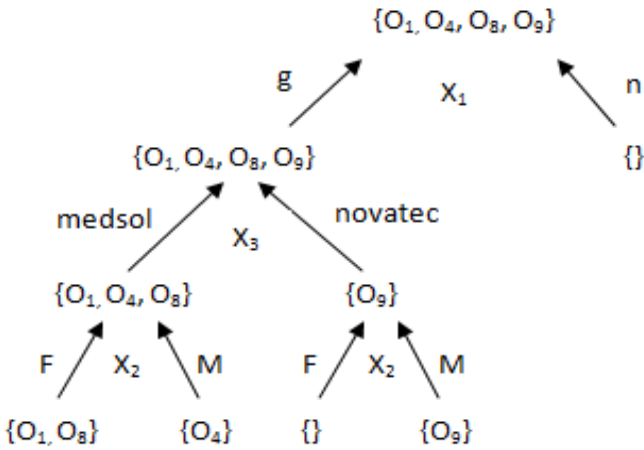
Para la creación de los árboles de decisión se deben crear tantos árboles como respuestas tenga el rasgo objetivo. En este caso como el valor del rasgo objetivo, ram principal, solo toma dos posibles valores, solo es necesario la creación de dos árboles de decisión  $Y_{11} = \text{Cefalea}$  y  $Y_{12} = \text{Vértigo}$ .

# Capítulo 2: Propuesta de solución del Sistema Basado en Casos

1. Se agrupan los casos en dependencia de la decisión, estos constituirán la raíz del árbol.
2. Los nodos hijos de la raíz van condicionados a las respuestas que puedan tomar los rasgos más asociados. Los niveles de cada árbol comienzan por el valor de asociación. Al ser el grupo etario ( $X_3$ ) el rasgo más asociado, los nodos hijos de dicha raíz se separan en dos grupos debido a que poseen dos valores:
  - Aquellos casos que toman valor “g” para  $X_3$
  - Aquellos casos que toman valor “n” para  $X_3$

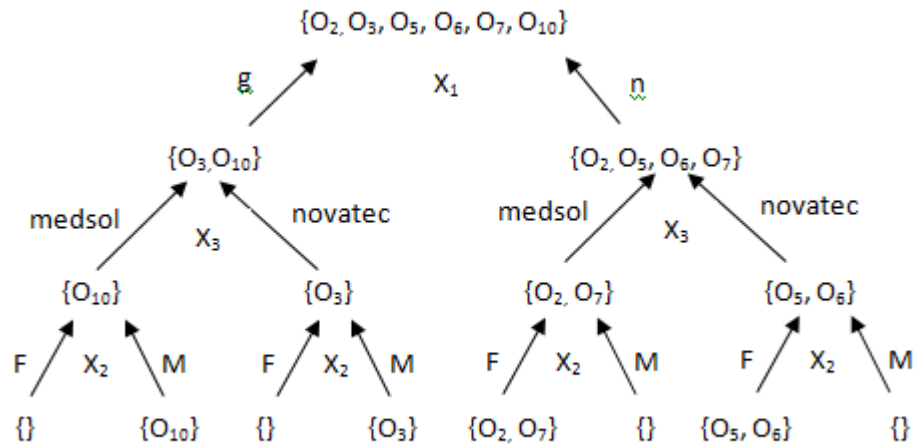
De esta forma quedan conformados los nodos hijos de la raíz, así como el nodo raíz del subárbol que sigue.

3. Se repite el proceso hasta que un nodo hijo no pueda ser separado por ningún otro rasgo, en ese caso se le considera hoja. A continuación se muestra los árboles de decisión asociados a  $Y_{11}$  y  $Y_{12}$ .



**Figura 10.** Árbol de decisión asociado al valor de la decisión  $Y_{11}$  (elaboración propia)





**Figura 11.** Árbol de decisión asociado al valor de la decisión Y<sub>12</sub> (elaboración propia)

**Séptimo paso:** Después de representar los rasgos asociados en los árboles de decisión se procede al cálculo de las primeras probabilidades de:

- Los rasgos no asociados fórmula 4
- Los rasgos asociados fórmula 5
- Los rasgos objetivos fórmula 6

**La probabilidad de los rasgos no asociados** se calcula, agrupando los casos atendiendo a los valores que tomen dichos rasgos, dividiendo los grupos formados entre el total de casos. En este caso no existen rasgos no asociados, pero de existir quedaría de la siguiente forma:

- Valor 1 = {O<sub>x</sub>, O<sub>x</sub>, O<sub>x</sub>, O<sub>x</sub>, O<sub>x</sub>}
- Valor 2 = {O<sub>x</sub>, O<sub>x</sub>, O<sub>x</sub>, O<sub>x</sub>, O<sub>x</sub>}

La cardinalidad de cada conjunto entre la unión de todos los conjuntos formados, para ambos casos sería 5/10.

**La probabilidad de los rasgos asociados** se calcula teniendo en cuenta la cardinalidad del conjunto hijo entre la cardinalidad del conjunto padre y así recursivamente. Para el conjunto de casos {O<sub>2</sub>, O<sub>7</sub>}, la

## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

cardinalidad sería dos y la cardinalidad de su raíz  $\{O_2, O_7\}$  es dos, la primera probabilidad para este conjunto de rasgos sería  $2/2$ .

**La probabilidad de los rasgos objetivos**, se aplica para cada caso y se determina como la razón entre la cardinalidad del conjunto de las hojas y la cardinalidad de la unión de dicho conjunto con el conjunto las hojas correspondientes al resto de los árboles formados. La razón obtenida es multiplicada por el producto de las probabilidades de los rasgos no asociados a dicho caso.

Por ejemplo, para los dos árboles formados anteriormente la probabilidad de los rasgos objetivos es:

Para el conjunto de casos  $\{O_3\}$  la cardinalidad es uno y la cardinalidad de la unión con el resto de las hojas de los demás árboles  $\{O_3, O_9\}$  es dos, como la probabilidad de los rasgos no asociados se comporta de igual manera para estos casos, basta con multiplicar el resultado de la razón por cero.

**Octavo paso:** cálculo de la segunda probabilidad

Para el cálculo de la segunda probabilidad de los rasgos predictivos se definen tres métodos:

- **Criterio de expertos**

Se refiere al índice de experticia de cada especialista sobre cada uno de los rasgos predictivos fórmula 8.

Para el rasgo  $X_1$  sería:

Suponiéndose que existen tres expertos  $a_1, a_2, a_3$  con grado de experticia  $c_1 = 0,8, c_2 = 0,9, c_3 = 0,8$  respectivamente. Los resultados de sus votaciones se muestran en la Tabla 9.

$$P_2 = \frac{0,8 * 1}{2,5} + \frac{0,9 * 0}{2,5} + \frac{0,8 * 1}{2,5}$$

**Fórmula 18:** Cálculo de la segunda probabilidad por criterio de expertos

**Tabla 9:** Resultados de las votaciones de los expertos (elaboración propia)

Asignación de valores	Expertos
$X_1(O_1) = g$	$a_1, a_3$
$X_1(O_2) = n$	$a_2, a_3$
$X_1(O_3) = n$	$a_1, a_2, a_3$
$X_1(O_4) = n$	$a_1, a_3$
$X_1(O_5) = g$	$a_1, a_2, a_3$
$X_1(O_6) = g$	$a_1, a_2, a_3$
$X_1(O_7) = g$	$a_1, a_2, a_3$

## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

$X_1(O_8) = n$	$a_1, a_2, a_3$
$X_1(O_9) = n$	$a_1, a_2, a_3$
$X_1(O_{10}) = g$	$a_1, a_2, a_3$

- **Certeza de instrumento de medición**

Se refiere al valor condicionado por un instrumento de medición utilizada en la fórmula 9.

Para el rasgo  $X_n$  sería:

Suponiéndose que en los casos  $O_i$  y  $O_j$  la temperatura se mide con un termómetro cuya certidumbre es uno y en los restantes casos con un termómetro cuya certidumbre es 0,8. Ambos valores constituyen la segunda probabilidad para dicho rasgo.

- **Fuente confiable**

Absoluta certeza por parte del especialista (la certeza es igual a 1) fórmula 7.

Para el cálculo de la segunda probabilidad de los rasgos objetivos fórmula 10, sería la sumatoria de todas las segundas probabilidades de cada rasgo predictivo de cada caso entre el número total de rasgos predictivos.

**Noveno paso:** cálculo de la certidumbre total asociada a los rasgos predictivos y objetivos

Después de tener ambas probabilidades calculadas, se procede a determinar el valor de la certidumbre total, se toman los valores de la primera probabilidad y los de la segunda. Este cálculo viene dado por la fórmula 11.

**Décimo paso:** cálculo de la semejanza a partir de un nuevo caso

Después de obtener las certidumbres totales, tanto para los rasgos predictivos como objetivos, se procede a recuperar los casos semejantes. Para esto se utiliza de un nuevo caso.

**Tabla 10:** Valores con sus certidumbres asociadas para el problema a resolver (elaboración propia)

Caso	$X_1$ Grupo Etario	$\xi_1(O_t)$	Sexo	$\xi_2(O_t)$	Fabricante	$\xi_3(O_t)$
$O_0$	n	0.82	F	1	MedSol	0.8

Para determinar la certidumbre total de cada rasgo predictivo en el nuevo caso, se toma como dicha probabilidad, el proceso de segunda probabilidad descrita en noveno paso.

## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

Para determinar los casos más semejantes, se tomó en consideración los casos correspondientes a la hoja dos de cada árbol, debido a que son los obtenidos al realizar una búsqueda rigiéndose por los valores de los rasgos predictivos. Luego se consideran los casos  $O_4$ ,  $O_6$  y  $O_8$  los más semejante, al aplicar la función de semejanza con el nuevo caso. Así se reducen ostensiblemente el número de comparaciones que se tiene que realizar.

Para obtener la semejanza, se divide el proceso en dos etapas:

- **Etapla 1: Función de similitud**

El cálculo de la similitud se obtiene mediante la fórmula 13 y se explica como la suma de las certidumbres totales calculadas del caso recuperado y el nuevo caso entre dos si; solo si ambas certidumbres son mayores que el umbral de semejanza  $\varepsilon = 0.8$ , siendo este valor un determinante de cuán exquisito será el sistema a la hora de determinar la solución. (Pérez, y otros, 2002 pág. 115)

- **Etapla 2: Función de semejanza**

El resultado de la semejanza se obtiene mediante la fórmula 12 y se explica como la sumatoria de los pesos obtenidos de cada rasgo por la función de similitud entre los rasgos; entre la sumatoria de todos los pesos.

Para los casos recuperados los valores de semejanzas se muestran a continuación:

$$\beta(O_0O_4) = 0.6, \quad \beta(O_0O_6) = 0.4, \quad \beta(O_0, O_8)$$

**Onceno paso:** Eliminación de la inconsistencia

Los casos más semejantes son  $O_4$  y  $O_8$ , que como puede apreciarse tienen los mismos valores para los rasgos predictivos y valores diferentes para el rasgo objetivo. Esta situación se denomina inconsistencia y es una de las problemáticas actuales de los Sistemas Basados en el Conocimiento. Para seleccionar el caso óptimo entre los casos recuperados y resolver el problema de la inconsistencia, se toma  $\gamma = 0,5$ , donde  $\gamma$  es un parámetro que se selecciona según el criterio del experto. A medida que  $\gamma$  tiende a uno, significa que se le está dando mayor importancia a la semejanza que a la certidumbre de la solución. (Pérez, y otros, 2002 pág. 115)

## Capítulo 2: Propuesta de solución del Sistema Basado en Casos

---

De acuerdo a la fórmula 15 entonces queda:

$$\mu(O_4) = 0.6 \quad \mu(O_8) = 0.8$$

El caso óptimo para retornar es el caso  $O_8$ , luego la solución ante el nuevo caso, es que podrá ocurrir una reacción adversa de Cefalea con un 80% de certeza.

### Capítulo 3: Validación del Sistema Basado en Casos

En el presente capítulo, en aras de realizar la validación del Sistema Basado en Casos, se presenta la interacción del especialista con el sistema; mostrando las interfaces de usuario. Además se hace uso del principio filosófico de razonamiento, conocido como principio de parsimonia o simplicidad, o Navaja de Ockham.

#### 3.1 Interfaz de usuario

El sistema incluye dentro de sus elementos fundamentales el componente de interfaz de usuario, el cual se estructura de las siguientes secciones:

- Área para la entrada de datos
- Área de resultados
- Área para ver detalles del resultado

A continuación se describen cada una de las secciones antes mencionadas.

- **Área para la entrada de datos**

Esta sección constituye el área donde el especialista, introducirá los datos relacionados con la prescripción a realizar al paciente en consulta y que serán necesarios para que el sistema realice la predicción.

The screenshot shows a web form titled "Introducir nuevo caso" with a blue arrow icon. The form contains six dropdown menus arranged in a 2x3 grid. The labels for the dropdowns are: "Medicamento sospechoso:", "Motivo prescripción:", "Antecedente patológico:", "Grupo etario:", "Sexo:", and "Fabricante:". Each dropdown menu has a light blue border and a small downward arrow on the right side. The text "<<Seleccione>>" is visible in each dropdown. At the bottom right of the form, there are two buttons: "Aceptar" and "Cancelar", both with light blue backgrounds and dark text.

Figura 12. Área para la entrada de datos

## Capítulo 3: Validación del Sistema Basado en Casos

- **Área de resultados**

En esta sección, el sistema muestra los resultados luego de realizada la predicción. Se presentan los diez casos más semejantes, siendo el primero el de mayor semejanza al caso conformado durante la entrada de los datos en el área para la entrada de datos; en aras de brindarle al especialista un margen de análisis que le facilite según su experiencia realizar la prescripción. Cada uno de los resultados está conformado por los principales datos referentes a los casos almacenados en la Base de Casos, siendo estos: medicamento sospechoso, motivo de prescripción, antecedente patológico, grupo etario, sexo, fabricante, ram principal y certeza.



Medicamento sospechoso	Motivo prescripción	Antecedente patológico	Grupo etario	Sexo	Fabricante	Ram principal	Certeza	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		a	F	Novatec	VOMITOS	100 %	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		a	F	Novatec	ABDOMEN, DOLOR	100 %	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		a	F	Novatec	ABDOMEN, DOLOR	100 %	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		g	F	Novatec	EPIGASTRICO, DO...	89 %	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		g	F	Novatec	EPIGASTRICO, DO...	89 %	

Figura 13. Área de resultados

- **Área para ver detalles del resultado**

Mediante esta área, el especialista podrá observar otros datos referentes a los resultados obtenidos luego de realizada la predicción. Para acceder a ellos solo deberá accionar el icono que aparece al final de cada resultado.



Medicamento sospechoso	Motivo prescripción	Antecedente patológico	Grupo etario	Sexo	Fabricante	Ram principal	Certeza	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		a	F	Novatec	VOMITOS	100 %	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		a	F	Novatec	ABDOMEN, DOLOR	100 %	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		a	F	Novatec	ABDOMEN, DOLOR	100 %	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		g	F	Novatec	EPIGASTRICO, DO...	89 %	
AZITROMICINA 250 MG CAPSULAS	Ira Baja		g	F	Novatec	EPIGASTRICO, DO...	89 %	

Figura 14. Ícono para ver detalles del resultado

The screenshot shows a window titled "Detalles del caso" with three expandable sections:

- Otras reacciones adversas:** Three input fields labeled "Reaccion adversa 1:", "Reaccion adversa 2:", and "Reaccion adversa 3:". The values entered are "NAUSEAS", "DIARREA", and "FATIGA" respectively.
- Otros farmacos:** Four input fields labeled "Farmaco 1:", "Farmaco 2:", "Farmaco 3:", and "Farmaco 4:". The values entered are "Dipirona", "Imefasma", and two empty fields.
- Otros datos:** Two input fields labeled "Severidad de la reacción:" and "Desenlace:". The values entered are "Moderada" and "Vivo".

A "Cerrar" button is located at the bottom right of the window.

**Figura 15.** Área ver detalles de resultado

El diseño de cada pantalla sigue las pautas y normas definidas por el Departamento de Sistemas de Apoyo a la Salud para el desarrollo del proyecto Synta.

### 3.2 Principio de la parsimonia o Navaja de Ockham

La navaja de Ockham proviene del filósofo inglés William Ockham, quien aplicó la idea de navaja para eliminar los supuestos innecesarios de una teoría. Esta posee su basamento en que *“ante dos o más explicaciones a un mismo hecho, la explicación más simple suele ser la más probable, pero no necesariamente verdadera”*. (La Guía, 2011)

El basamento antes planteado se puede explicar de la siguiente forma: si existen dos o más explicaciones que estén en igualdad de condiciones, no existe la necesidad de tener en cuenta una explicación complicada si existe una más simple. Esto no quiere decir que la explicación más simple sea la más correcta, sino que existen más probabilidades que sea cierta y que es preferible elegirla hasta que haya razones bien fundamentadas para adoptar una alternativa más compleja.

### 3.3 Aplicación del principio de la parsimonia o Navaja de Ockham en la validación de la solución

A continuación se presentan un conjunto de interrogantes que tributarán a un mejor entendimiento del principio utilizado:



¿Qué sucedería si al momento del médico introducir los datos necesarios para realizar la predicción del sistema propuesto los datos entrados coincidiesen en un 100% con los de algún caso almacenado en Base de Casos? Pues la semejanza entre el caso nuevo y uno almacenado deberá ser del 100 por ciento y por ende la certeza de predicción sería de igual por ciento. Por otra parte ¿Qué sucedería con la certeza de predicción cuando al introducir primeramente un caso nuevo, donde variase el valor de uno de los rasgos predictores, por ejemplo: el grupo etario y en un segundo momento variase el valor de los rasgos grupo etario y sexo con respecto a alguno de los almacenados en la Base de Casos? Pues la explicación más sencilla dice que la certeza de predicción resultante ante el primer resultado debe ser menor que cien, pero mayor que la obtenida ante el segundo resultado. Siguiendo el principio descrito anteriormente se procedió a la validación del SBC, obteniéndose los resultados que se mostrarán en el presente capítulo.

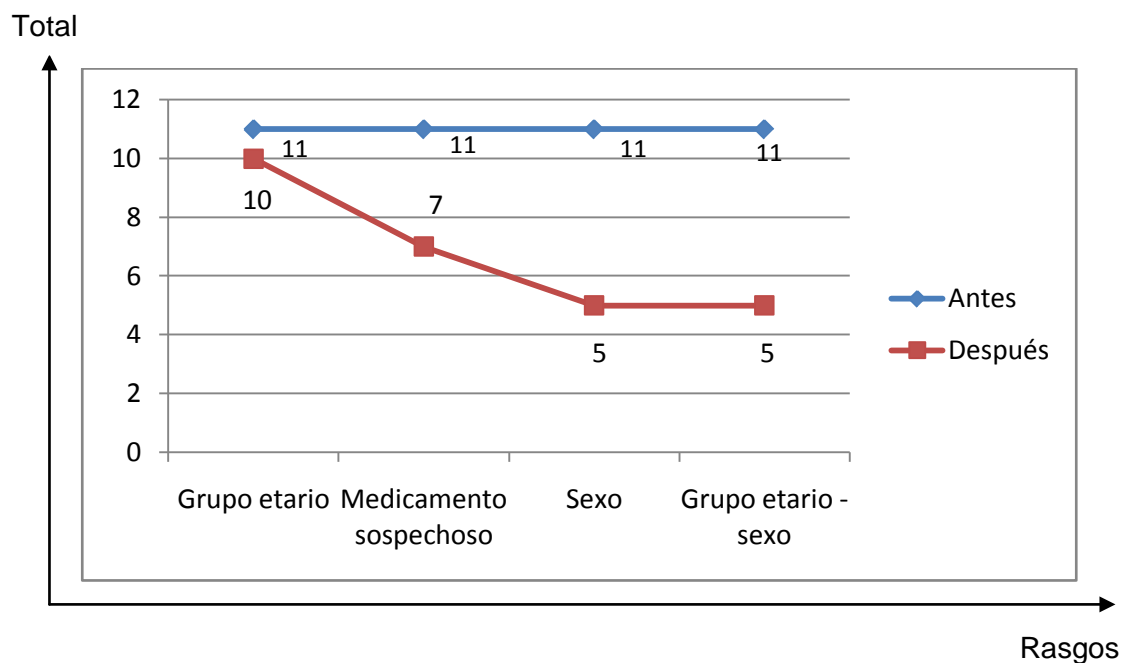
El SBC posee una Base de Casos de 7927 registros de RAM (casos), los cuales han sido registrados por el departamento de Farmacovigilancia de La Habana durante el año 2010. Para la realización de los casos de prueba cuyos resultados validarían la solución se decide tener en cuenta principalmente los casos cuyo desenlace fuese *mortal* y *grave* debido a que con la prevención de estos contribuiría a una disminución en la severidad de las RAM en los pacientes así como de los costos asociados al tener que tratar las RAM.

En un primer momento se tomaron todos los casos que tuvieron como desenlace *fallecido*, obteniéndose once resultados, a los cuales se les aplicó el principio de la navaja de Ockham antes descrito a través de los pasos que se describen a continuación:

1. Obtención de los casos almacenados cuyo desenlace fuese *fallecido*.
2. Introducir los casos obtenidos en el sistema variando el valor del rasgo predictor *grupo etario* por uno de mayor peso o relevancia.
3. Introducir los casos obtenidos en el sistema variando el valor del rasgo predictor *medicamento sospechoso*.
4. Introducir los casos obtenidos en el sistema variando el valor del rasgo predictor *sexo*.
5. Introducir nuevamente los casos obtenidos en el sistema variando el valor de los rasgos predictores *grupo etario* y *sexo*.

## Capítulo 3: Validación del Sistema Basado en Casos

Tras la realización de los pasos anteriormente descritos se obtienen resultados con una certeza de predicción promedio de 72,04%, los cuales se muestran en la siguiente gráfica:

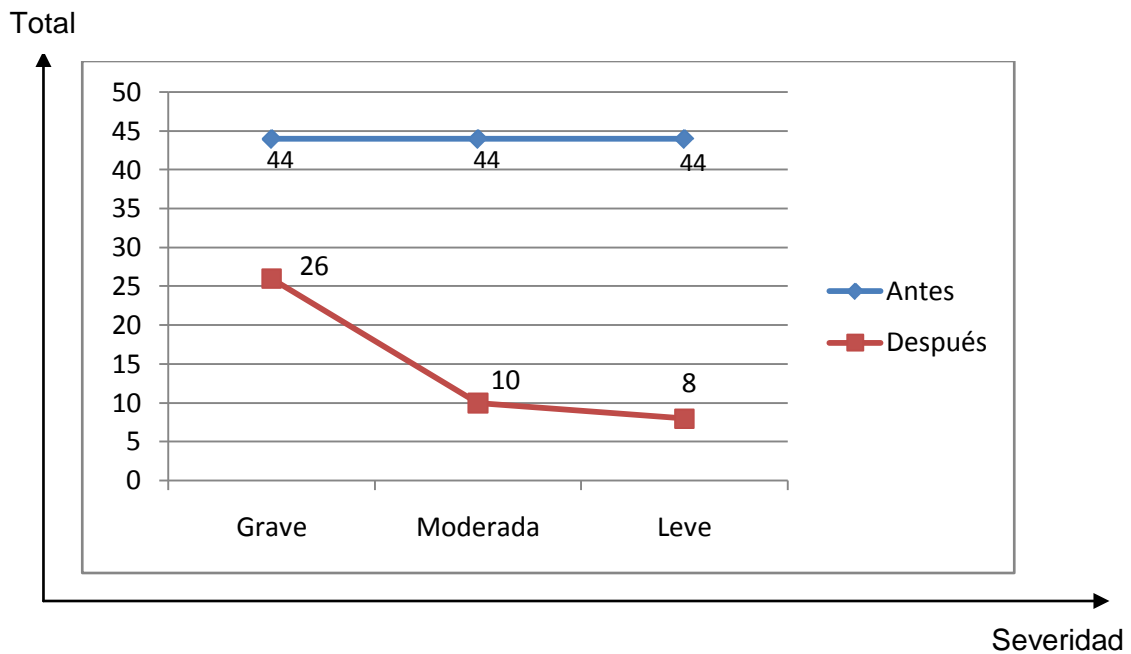


**Figura 16.** Resultados obtenidos para los casos con desenlace fallecidos (elaboración propia)

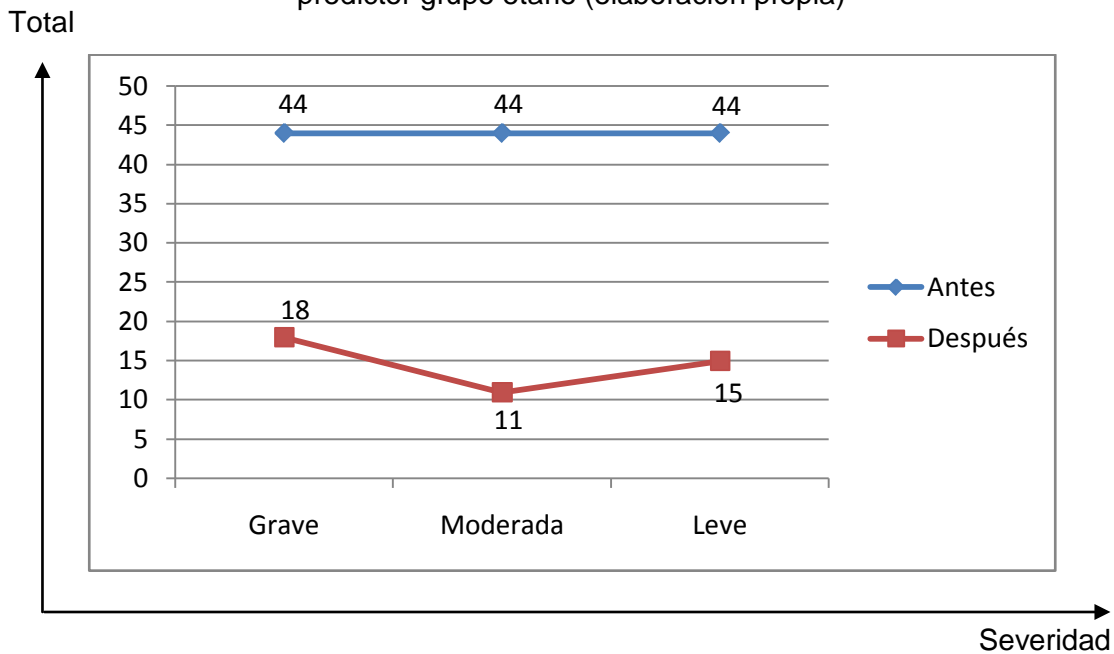
En un segundo momento se procedió a la validación del SBC ante la severidad grave de las RAM, con este fin se definieron los pasos descritos a continuación:

1. Obtención de los casos almacenados cuya severidad fuese grave
2. Introducir los casos obtenidos en el sistema variando el valor del rasgo predictor grupo etario
3. Introducir los casos obtenidos en el sistema variando el valor del rasgo predictor sexo

Obteniéndose el resultado, con una certeza de predicción promedio de 76,54% variando los valores del rasgo predictor grupo etario, mientras que para el rasgo predictor sexo, la certeza de predicción promedio fue de un 66,72%. Los resultados se muestran en las siguientes figuras:



**Figura 17.** Resultados obtenidos para los casos con severidad grave variándole el valor del rasgo predictor grupo etario (elaboración propia)



**Figura 18.** Resultados obtenidos para los casos con severidad grave variándole el valor del rasgo predictor grupo etario (elaboración propia)

La Figura 16 representa los resultados obtenidos luego de la predicción, para los casos cuyo desenlace es fatal. Para graficar dichos resultados se variaron los valores de los rasgos predictores: grupo etario, medicamento sospechoso, sexo y la combinación grupo etario-sexo respectivamente. Mientras que las Figuras 17 y 18 muestran el comportamiento de la severidad ante el cambio del valor de los rasgos predictores grupo etario y sexo respectivamente.

En todos los casos de prueba, se variaron los valores de sexo, grupo etario y medicamento sospechoso por ser los valores de mayor peso en el momento de identificar qué RAM le podría causar a un paciente al ingerir un nuevo medicamento. Además demostraron que puede existir una disminución de la ocurrencia de las RAM y severidad de las mismas (respecto al nivel, pueden bajar de graves a moderadas, o a leves), si se utiliza el SBC.

### Conclusiones

Con la realización del presente trabajo de diploma se ha cumplido con el objetivo general propuesto, así como con las tareas de la investigación obteniéndose las siguientes conclusiones:

- El estudio de los principales Sistemas Expertos encontrados evidenció, que no poseen las características necesarias para posibilitar la predicción de las reacciones adversas a los medicamentos desde la consulta médica.
- Analizada la estructura y clasificación de los Sistemas Expertos, se determinó que los Sistemas Basado en Casos constituían una solución óptima para realizar la predicción de las Reacciones Adversas a Medicamentos, teniendo en cuenta la información con lo que se contaba.
- Con la utilización del conjunto de Testores Típicos se identificaron de un total de setenta y seis rasgos, seis constituían los rasgos predictores a utilizar en la Máquina de Inferencia.
- El estudio de las principales técnicas utilizadas para la implementación de la Máquina de Inferencia, evidenció que las utilizadas en la investigación constituían las más adecuadas a utilizar en condiciones de incertidumbre.
- Con la obtención de la frecuencia esperada en el momento de la obtención de la decisión se reduce el problema de la inconsistencia en los Sistemas Basados en Casos.
- Se asimilaron las herramientas y tecnologías propuestas por el departamento Sistemas de Apoyo a la Salud (SAS), lo que permitió obtener un SBC que permite la predicción de las RAM en los pacientes en la consulta médica.

### Recomendaciones

Al finalizar la presente investigación se recomienda:

- Implementar un Sistema Basado en Reglas e incorporar al Sistema Basado en Casos, que permita al médico en el momento de la consulta realizar la comparación entre el medicamento a prescribir y otros medicamentos que consume el paciente. Indicando al médico si el medicamento prescrito puede causar interacción medicamentosa.
- Realizar interfaces en el módulo de Reacciones Adversas a Medicamentos, que permita al experto poder variar el peso de los rasgos predictores utilizados en el motor de inferencia.
- Implementar la Técnica Chi-Square Automatic Interaction Detection (CHAID) en el Sistema Basado en Casos, con el objetivo de agilizar el acceso a los datos en la Base de Casos mejorando el tiempo de respuesta del sistema.
- Tener en cuenta el nomenclador Anatómico Terapéutico Químico (ATC) del medicamento sospechoso en la recuperación de los casos más semejante, otorgando mayor prioridad a los casos, cuyo medicamento sospechoso pertenezca al mismo grupo del nomenclador ATC.
- Incorporar una interfaz al Módulo de Reacciones Adversas a Medicamentos, donde se explique el funcionamiento del Sistema Basado en Casos en lenguaje natural, al experto, logrando este un mayor entendimiento sobre el funcionamiento del sistema y aumentando de esta manera la simpatía para su uso.

## Referencias bibliográficas

1. (Agudo, 2001 p. 37) **Agudo, B. 2001.** *Introducción al razonamiento basado en casos. Departamento de sistemas informaticos y programacion.* Espana : Universidad complutense de madrid, 2001.
2. (Agudo, 2009 pág. 6) **Sistema experto ONCOCIN para el tratamiento del cáncer. Agudo, David Del Valle. 2009.** 2009, pág. 6. 100025228.
3. (ANMAT, 2009 pág. 110) **ANMAT. 2009.** *GUÍA DE BUENAS PRÁCTICAS DE FARMACOVIGILANCIA.* Ciudad Autónoma de Buenos Aires : Ministerio de Salud, 2009. pág. 110. s.n.
4. (Avanzadas, 2007 p. 7) **Avanzadas, Grupo de Bases de Datos. 2007.** *Concepto y Objetivos de las Bases de Datos.* Madrid : s.n., 2007.
5. (Bello, 2002, pág. 100) **Aplicaciones de la Inteligencia Artificial. Pérez, Rafael E. Bello. 2002.** Guadalajara, Jalisco : Ediciones de la Noche, 28 de 02 de 2002. 970-27-0177-5.
6. (Budnitz, et al., 2012 pág. 11) **Emergency Hospitalizations for Adverse Drug Events in Older Americans. Budnitz, Daniel S., Lovegrove, Maribeth C. y Richards, Chesley L. 2012.** Massachussets : The new England Journal of Medicine, 2 de 11 de 2012, pág. 11. 365:2002-12.
7. (CCS, 2012) **CCS. 2012.** Libroweb.es. *¿Qué es CSS?* [En línea] 2012. [Citado el: 28 de 03 de 2013.] <http://www.librosweb.es/css/capitulo1.html>
8. (Ciberaula, 2010) **Ciberaula. 2010.** Una Introducción a APACHE. [En línea] 2010. [Citado el: 28 de 03 de 2013.]  
  
[http://linux.ciberaula.com/articulo/linux\\_apache\\_intro](http://linux.ciberaula.com/articulo/linux_apache_intro)
9. (Farmacovigilancia, 2010 pág. 12) **Farmacovigilancia, Unidad Coordinadora Nacional de. 2010.** *Comunicación a profesionales sanitarios sobre identificación de señales en reacciones adversas de baja frecuencia.* Departamento de Farmacoepidemiología, Ministerio de Salud Pública (MINSAP). La Habana : s.n., 2010. pág. 12, Comunicación.
10. (Javascript, 2009) **Javascript Ya. 2009.** Javascript Ya. [En línea] 2009. [Citado el: 28 de 03 de 2013.] <http://www.javascriptya.com.ar>
11. (La Guía, 2011) **La Guía. 2011.** La Guía. *Filosofía.* [En línea] 11 de 08 de 2011.

- [Citado el: 06 de 05 de 2013.] <http://filosofia.laguia2000.com/ciencia-y-filosofia/la-navaja-de-ockham>
12. (Laporte, y otros, 2007  
pág. 247) **Laporte, JR y Tognoni, G. 2007.** *Principios de epidemiología del medicamento.* s.l. : Masson Salvat Medicina, 2007. pág. 247.
  13. (Lio, 1998) **Sistema Inteligente de Selección de Información SISI. Lio, Dr. Daniel Galvez. 1998.** Universidad Central “Martha Abreu” de Las Villas : Departamento de ciencias de computación. : s.n., 1998
  14. (López, 2010) **López, MSc. Luisa Velásquez. 2010.** Estructura de un sistema experto. [En línea] Instituto de Investigaciones en Informática, 19 de 08 de 2010. [Citado el: 06 de 12 de 2012.] [http://iii.informatica.edu.bo/index.php?option=com\\_content&view=article&id=106:eduse&catid=43:revista2007&Itemid=60](http://iii.informatica.edu.bo/index.php?option=com_content&view=article&id=106:eduse&catid=43:revista2007&Itemid=60)
  15. (Lorenzo, 2000 pág. 58) **Lorenzo Castro, Fernando. 2000.** *MODELOS DE DATOS CONCEPTOS Y CLASIFICACIÓN.* 2000. pág. 58.
  16. (Márquez, 2010) **Márquez, Samper y José, Juan. 2010.** Red Científica. [En línea] 2010. [Citado el: 20 de 01 de 2013.] <http://www.redcientifica.com/doc/doc199908210001.html>. 1579-0223.
  17. (Martínez, et al., 2008,  
pág. 63) **Martínez Sánchez Natalia, Gheisa Ferreira Lorenzo, María M. García Lorenzo, Zenaida García Valdivia. 2008.** *El Razonamiento Basado en Casos en el ámbito de la Enseñanza/Aprendizaje.* s.l. : Revista de Informática Educativa y Medios Audiovisuales, 2008. 1667-8338.
  18. (MySql, 2011) **MySql. 2011.** MySql: The world's most popular open source database. [En línea] 2011. [Citado el: 03 de 28 de 2013.] <http://dev.mysql.com/doc/refman/5.0/es/>
  19. (Netbeans, 2010) **Netbeans. 2010.** Equipo de desarrollo, desarrolloweb.com. [En línea] 22 de 06 de 2010. [Citado el: 28 de 03 de 2013.] <http://www.desarrolloweb.com/actualidad/netbeans-6-9-final-liberado-3613.html>
  20. (NDS, 2009) **NDS. 2009.** Neurology Diagnosis System. [En línea] NDS, 06 de 04 de 2009. [Citado el: 14 de 05 de 2013.] <http://ireasoning.sourceforge.net/>
  21. (Ochoa, 2004 p. 37) **Ochoa Somuano, Jorge. 2004.** *Reduccion de Dimensiones en Base de Datos usando el enfoque Logico Combinatorio.* 2004. 968-878-205-X.



22. (OMS, 2012) **OMS. 2012.** Organización mundial de la Salud. *Farmacovigilancia*. [En línea] 09 de 11 de 2012. [Citado el: 26 de 11 de 2012.] <http://apps.who.int/medicinedocs/en/d/Js6166s/4.html>
23. (OMS, 2013a) **OMS. 2013.** Organización Mundial de la Salud. *OMS*. [En línea] 2013. [Citado el: 08 de 01 de 2013.] <http://www.who.int/mediacentre/factsheets/fs293/es/index.html>
24. (OMS, 2013) **OMS. 2013.** Organización Mundial de la Salud. [En línea] 2013. [Citado el: 08 de 01 de 2013.] <http://www.who.int/about/es/>
25. (PHP, 2012) **PHP. 2012.** Manual de PHP. [En línea] 08 de 06 de 2012. [Citado el: 28 de 03 de 2013.] <http://www.php.net/manual/es/index.php>
26. (Pérez, y otros, 2002 pág. 107) **Un SBC para la toma de decisiones en condiciones de incertidumbre. Pérez, Rafael E. Bello, Gutiérrez Martínez, Iliana y Tellería Rodríguez, Andrés. 2002.** 2, Santa Clara : s.n., 2002, Vol. 23.
27. (Pons, A., 2003) **LEX: A new algorithm for the calculus of all typical. Pons, A. 2003.** Cuba : Revista Ciencias, 2003.
28. (Prieto, 2011 pág. 24) **Prieto, Osvaldo Caseres. 2011.** *Sistema Inteligente para el Diagnóstico de Hipertensión Arterial*. Habana : Universidad de las Ciencias Informáticas, 2011.
29. (Quality, 2001) **Quality, Agency for Healthcare Research and. 2001.** Agency for Healthcare Research and Quality. *Reducing and Preventing Adverse Drug Events To Decrease Hospital Costs*. [En línea] 03 de 2001. [Citado el: 25 de 10 de 2012.] <http://www.ahrq.gov/qual/aderia/aderia.htm>. 01-0020
30. (Quintana, y otros, 2004 pág. 53) **Quintana, Mariledy y Marrero, Mercedes. 2004.** *La Gestión Ambiental dentro de la Planificación Estratégica de la Ciudad de Matanzas*. Matanzas : s.n., 2004. pág. 53.
31. (Rodríguez, 2003 p. 58) **Rodríguez, Yalila Arean. 2003.** *Una Herramienta de búsqueda inteligente en bases de datos utilizando técnicas de RBC*. Habana : s.n., 2003.
32. (Rodríguez, 2010 pág. 209) **Sistema experto basado en casos para el diagnóstico de la hipertensión arterial. Rodríguez, Santiago Cuadrado. 2010.** 60, Colombia : Revista de Ingeniería de la Universidad de Antioquia, 15 de 11 de 2010. 0120-6230
33. (Rodríguez, 2007) **Fernández Rodríguez, Ricardo. 2007.** Centro de Cibernética

- Aplicada a la Medicina. *Centro de Cibernética Aplicada a la Medicina (CECAM)*. [En línea] 2007. [Citado el: 10 de 06 de 2013.] [http://www.rcim.sld.cu/revista\\_12/articulos\\_htm/razonamiento.htm](http://www.rcim.sld.cu/revista_12/articulos_htm/razonamiento.htm). 1684-185.
34. (Rodríguez, 2010 pág. 209) *Sistema experto basado en casos para el diagnóstico de la hipertensión arterial. Rodríguez, Santiago Cuadrado. 2010.* 60, Colombia : Revita de Ingeniería de la Universidad de Antioquia, 15 de 11 de 2010. 0120-6230
35. (Santealla, 2005 p. 8) **Santealla Vallejo, Juan. 2005.** *Un sistema experto: MYCIN.* Universidad Carlos III. Madrid : s.n., 2005. pág. 8.
36. (Schmieg, 2008) **Schmieg, Sebastian. 2008.** *Fundamentos Informaticos.* [En línea] 02 de 02 de 2008. [Citado el: 20 de 01 de 2013.] <http://fundamentosinformaticosjl.wordpress.com>
37. (Sencha, 2011) **Sencha. 2011.** *Sencha.* [En línea] 2011. [Citado el: 28 de 03 de 2013.] <http://www.sencha.com/products/extjs>
38. (SQL Manage, 2013) **Manager, Sql. 2013.** *Sql Manager. EMS SQL Manager for MySQL.* [En línea] EMS, 2013. [Citado el: 28 de 03 de 2013.] <http://www.sqlmanager.net/es/products/mysql/manager>
39. (Uppsala, et al., 2001 p. 32) **Uppsala Monitoring Centre, Centre for International Drug Monitoring Stora Torget. 2001.** *Vigilancia de la Seguridad de los Medicamentos.* the Uppsala Monitoring Centre. Sweden : s.n., 2001. pág. 32. 91-631-1561-1.
40. (Vega, et al.,2000 pág. 8) *Razonamiento basado en casos aplicado al diagnóstico de generadores eléctricos. Vega, H. Octavio De la Torre y Lopez A., Ernesto. 2000.* Mexico : Boletín iie, Enero - Febrero de 2000, pág. 8.
41. (Zaninotto, y otros, 2008 pág. 11) **Zaninotto, François y Potencier, Fabien. 2008.** *Symfony la Guía Definitiva.* New York. : Appress, 2008. ISBN-13: 978-1590597866.

## Bibliografía

1. **Agudo, B. 2001.** *Introducción al razonamiento basado en casos. Departamento de sistemas informáticos y programación.* España : Universidad Complutense de Madrid, 2001.
2. **ANMAT. 2009.** *GUÍA DE BUENAS PRÁCTICAS DE FARMACOVIGILANCIA.* Ciudad Autónoma de Buenos Aires : Ministerio de Salud, 2009. pág. 110. s.n.
3. *Aplicaciones de la Inteligencia Artificial.* **Pérez, Rafael E. Bello. 2002.** Guadalajara, Jalisco : Ediciones de la Noche, 28 de 02 de 2002. 970-27-0177-5.
4. **Avanzadas, Grupo de Bases de Datos. 2007.** *Concepto y Objetivos de las Bases de Datos.* Madrid : s.n., 2007. s.n.
5. **CCS. 2012.** HTML.Net. *¿Qué es CSS?* [En línea] 2012. [Citado el: 28 de 03 de 2013.] <http://es.html.net/tutorials/css/lesson1.php>.
6. **Ciberaula. 2010.** Una Introducción a APACHE. [En línea] 2010. [Citado el: 28 de 03 de 2013.] [http://linux.ciberaula.com/articulo/linux\\_apache\\_intro](http://linux.ciberaula.com/articulo/linux_apache_intro).
7. *Conceptual Maps and Cases Based Reasoning: A perspective for the Intelligent Teaching-Learning Systems.* **Natalia, Martínez Sánchez, y otros. 2010.** 5, s.l. : IEEE LATIN AMERICA TRANSACTIONS, 2010, Vol. 8. s.n.
8. **de la Osa, José A. 2012.** Diario Granma. *Cubanos con reportes de reacciones adversas por antibióticos .* [En línea] ÓRGANO OFICIAL DEL COMITÉ CENTRAL DEL PCC, 18 de 09 de 2012. [Citado el: 20 de 09 de 2012.] <http://www.granma.cubaweb.cu/2012/09/18/nacional/artic02.html>. 258.
9. *El Razonamiento Basado en Casos en el ámbito de la enseñanza/aprendisaje.* **Martínez, Sánchez Natalia, y otros. 2008.** 10, Santa Clara : Revista de Informática Educativa y Medios Audiovisuales, noviembre de 2008, Revista de Informática Educativa y Medios Audiovisuales , Vol. 5, págs. 22-33. <http://laboratorios.fi.uba.ar/lie/Revista/articulos.htm>. ISSN 1667-8338.

10. *Emergency Hospitalizations for Adverse Drug Events in Older Americans*. **Budnitz, Daniel S., Lovegrove, Maribeth C. y Richards, Chesley L.** 2012. Massachusetts : The new England Journal of Medicine, 2 de 11 de 2012, pág. 11. 365:2002-12.
11. **Farmacovigilancia, Unidad Coordinadora Nacional de.** 2010. *Comunicación a profesionales sanitarios sobre identificación de señales en reacciones adversas de baja frecuencia*. Departamento de Farmacoepidemiología, Ministerio de Salud Pública (MINSAP). Ciudad de la Habana : s.n., 2010. pág. 12, Comunicación.
12. *Feature Selection using Typical Testors applied to Estimation of Stellar Parameters*. **Santos, José A., Carrasco, Ariel y Martínez, José F.** 2004. 1, Mexico : Computación y Sistemas, 2004, Vol. 18. ISSN 1405-5546.
13. **Fernández Rodríguez, Ricardo.** 2007. Centro de Cibernética Aplicada a la Medicina. *Centro de Cibernética Aplicada a la Medicina (CECAM)*. [En línea] 2007. [Citado el: 10 de 06 de 2013.] [http://www.rcim.sld.cu/revista\\_12/articulos\\_hm/razonamiento.htm](http://www.rcim.sld.cu/revista_12/articulos_hm/razonamiento.htm). 1684-185.
14. **Guerra-Gandón, Alejandro, Vega-Pons, Sandro y Ruiz-Shulcloper, José.** 2012. *Algoritmos de agrupamiento conceptuales: un estado del arte*. Ciudad de la Habana : Serie Azul, 2012. ISSN-2072-6287.
15. **Gutiérrez Martínez, Iliana.** 2001. *Un Modelo para la Toma de Decisiones en un Sistema Basado en Casos en condiciones de Incertidumbre*. santa Clara : s.n., 2001.
16. **HTML.** 2013. HTML.Net. *¿Qué es HTML?* [En línea] 2013. [Citado el: 13 de febrero de 2013.] <http://es.html.net/tutorials/html/lesson2.php>.
17. **Javascript.** 2009. Javascript. [En línea] 2009. [Citado el: 28 de 03 de 2013.] <http://www.javascriptya.com.ar>.
18. **KAUFMANN, M.** 2000. *Data Mining. Concepts and Techniques*. 2000.
19. **La Guía.** 2011. La Guía. *Filosofía*. [En línea] 11 de 08 de 2011. [Citado el: 06 de 05 de 2013.] <http://filosofia.laguia2000.com/ciencia-y-filosofia/la-navaja-de-ockham>.

20. **Laporte, JR y Tognoni, G. 2007.** *Principios de epidemiología del medicamento.* s.l. : Masson Salvat Medicina, 2007. pág. 247.
21. *LEX: A new algorithm for the calculus of all typical.* **Pons, A. 2003.** Cuba : Revista Ciencias, 2003.
22. **López, MSc. Luisa Velásquez. 2010.** Estructura de un sistema experto. [En línea] Instituto de Investigaciones en Informática, 19 de 08 de 2010. [Citado el: 06 de 12 de 2012.] [http://iii.informatica.edu.bo/index.php?option=com\\_content&view=article&id=106:eduse&catid=43:revista2007&Itemid=60](http://iii.informatica.edu.bo/index.php?option=com_content&view=article&id=106:eduse&catid=43:revista2007&Itemid=60).
23. **Lorenzo, Castro Fernando. 2000.** *MODELOS DE DATOS CONCEPTOS Y CLASIFICACIÓN.* 2000. pág. 58.
24. **Márquez, Juan José Samper. 2010.** Red Científica. [En línea] 2010. [Citado el: 20 de 01 de 2013.] <http://www.redcientifica.com/doc/doc199908210001.html>. 1579-0223.
25. **Márquez, Samper y José, Juan. 2010.** Red Científica. [En línea] 2010. [Citado el: 20 de 01 de 2013.] <http://www.redcientifica.com/doc/doc199908210001.html>. 1579-0223.
26. **MySQL. 2011.** MySQL: The world's most popular open source database. [En línea] 2011. [Citado el: 03 de 28 de 2013.] <http://dev.mysql.com/doc/refman/5.0/es/>.
27. **NDS. 2009.** Neurology Diagnosis System. [En línea] NDS, 06 de 04 de 2009. [Citado el: 14 de 05 de 2013.] <http://ireasoning.sourceforge.net/>.
28. **Netbeans. 2010.** Equipo de desarrollo, desarrolloweb.com. [En línea] 22 de 06 de 2010. [Citado el: 28 de 03 de 2013.] <http://www.desarrolloweb.com/actualidad/netbeans-6-9-final-liberado-3613.html>.
29. **Ochoa, Somuano Jorge. 2004.** *Reduccion de Dimensiones en Base de Datos usando el enfoque Logico Combinatorio.* 2004. 968-878-205-X.
30. **OMS. 2012.** Organizacion mundial de la Salud. *Farmacovigilancia.* [En línea] 09 de 11 de 2012. [Citado el: 26 de 11 de 2012.] <http://apps.who.int/medicinedocs/en/d/Js6166s/4.html>.

31. **OMS. 2013.** Organizacion Mundial de la Salud. OMS. [En línea] 2013. [Citado el: 08 de 01 de 2013.] <http://www.who.int/mediacentre/factsheets/fs293/es/index.html>.
32. **OMS. 2013a.** Organizacion Mundial de la Salud. [En línea] 2013a. [Citado el: 08 de 01 de 2013.] <http://www.who.int/about/es/>.
33. **Osa, José A. de la. 2012.** Granma. *Cubanos con reportes de reacciones adversas por antibióticos.* [En línea] 18 de 09 de 2012. [Citado el: 25 de 10 de 2012.] <http://www.granma.cubaweb.cu/2012/09/18/nacional/artic02.html>. 258.
34. **PHP. 2012.** Manual de PHP. [En línea] 08 de 06 de 2012. [Citado el: 28 de 03 de 2013.] <http://www.php.net/manual/es/index.php>.
35. **Prieto, Osvaldo Caseres. 2011.** *Sistema Inteligente para el Diagnóstico de Hipertensión Arterial.* Habana : Universidad de las Ciencias Informaticas, 2011. s.n.
36. *Principales resultados del sistema cubano de Farmacovigilancia en el año 2004.* **Debesa García, Francisco, y otros. 2005.** 3, Ciudad de la Habana : Revista Cubana de Farmacia, 2005, Vol. 39. ISSN 1561-2988.
37. **Quality, Agency for Healthcare Research and. 2001.** Agency for Healthcare Research and Quality. *Reducing and Preventing Adverse Drug Events To Decrease Hospital Costs.* [En línea] 03 de 2001. [Citado el: 25 de 10 de 2012.] <http://www.ahrq.gov/qual/aderia/aderia.htm>. 01-0020.
38. **Quintana, Mariledy y Marrero, Mercedes. 2004.** *La Gestión Ambiental dentro de la Planificación Estratégica de la Ciudad de Matanzas.* Matanzas : s.n., 2004. pág. 53.
39. *Razonamiento basado en casos aplicado al diagnóstico de generadores eléctricos.* **Vega, H. Octavio De la Torre y Lopez A., Ernesto. 2000.** Mexico : Boletín iie, Enero - Febrero de 2000, pág. 8.
40. **Rodriguez, Yalila Arean. 2003.** *Una Herramienta de busqueda inteligente en bases de datos utilizando tecnicas de RBC.* Habana : s.n., 2003.

41. **Santealla, Vallejo Juan. 2005.** *Un sistema experto: MYCIN*. Universidad Carlos III. Madrid : s.n., 2005. pág. 8.
42. **Santiesteban Alganza, Yovanis y Pons Porrata, Aurora. 2003.** *LEX: UN NUEVO ALGORITMO PARA EL CÁLCULO DE LOS TESTORES*. Dpto. de Computación, Universidad de Oriente. Santiago de Cuba : s.n., 2003. pág. 15.
43. **Schmieg, Sebastian. 2008.** Fundamentos Informaticos. [En línea] 02 de 02 de 2008. [Citado el: 20 de 01 de 2013.] <http://fundamentosinformaticosjl.wordpress.com/>.
44. **Sencha. 2011.** Sencha. [En línea] 2011. [Citado el: 28 de 03 de 2013.] <http://www.sencha.com/products/extjs>.
45. *Sistema de gestión de la promoción turística dirigido a la web semántica.* **Valdés Ramírez, Danilo. 2011.** 7, Ciego de Avila : Serie Científica de la Universidad de las Ciencias Informáticas, 2011, Vol. 4. s.n.
46. *Sistema experto basado en casos para el diagnóstico de la hipertensión arterial.* **Rodríguez, Santiago Cuadrado. 2010.** 60, Colombia : Revista de Ingeniería de la Universidad de Antioquia, 15 de 11 de 2010. 0120-6230.
47. *Sistema experto ONCOCIN para el tratamiento del cáncer.* **Agudo, David Del Valle. 2009.** 2009, pág. 6. 100025228.
48. *Sistema Experto para el diagnóstico médico de las enfermedades genéticas con dismorfias (SEGEDIS).* **Gutiérrez Rodríguez, Marianela, Barroso Rodríguez, Yadira y Bedoya Rusenko, Jorge. 2011.** 9, Ciudad de la Habana : Serie Científica de la Universidad de las Ciencias Informáticas, 2011, Vol. 4. s.n.
49. *Sistema Inteligente de Selección de Información SISI.* **Lio, Dr. Daniel Galvez. 1998.** Universidad Central "Martha Abreu" de Las Villas : Departamento de ciencias de computación. : s.n., 1998.
50. **SQL Manage. 2013.** Sql Manager. *EMS SQL Manager for MySQL*. [En línea] EMS, 2013. [Citado el: 28 de 03 de 2013.] <http://www.sqlmanager.net/es/products/mysql/manager>.

51. *Un SBC para la toma de decisiones en condiciones de incertidumbre*. **Pérez, Rafael E. Bello, Gutiérrez Martínez, Iliana y Tellería Rodríguez, Andrés. 2002.** 2, Santa Clara : s.n., 2002, Vol. 23.
52. *UNA APROXIMACIÓN MULTIMODAL AL DIAGNÓSTICO TEMPORAL MEDIANTE RAZONAMIENTO BASADO EN CASOS Y RAZONAMIENTO BASADO EN MODELOS. APLICACIONES EN MEDICINA*. **Juárez Herrero, José Manuel. 2007.** 036, Valencia, España : Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial., 2007, Vol. 11. 1137-3601.
53. **Uppsala, Monitoring Centre y Torget, Centre for International Drug Monitoring Stora. 2001.** *Vigilancia de la Seguridad de los Medicamentos*. the Uppsala Monitoring Centre. Sweden : s.n., 2001. pág. 32. 91-631-1561-1.
54. **Zaninotto, François y Potencier, Fabien. 2008.** *Symfony la Guía Definitiva*. New York : Appress, 2008. ISBN-13: 978-1590597866.



## Anexos

**Anexo 1:** Acumulado de informes de RAM recibidos por la OMS hasta 2004.

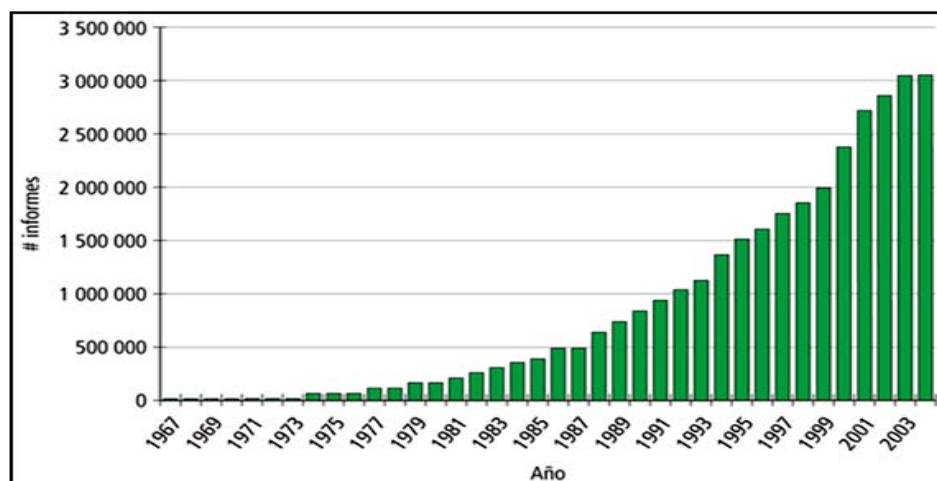


Figura 3 Número acumulado de informes recibidos por la OMS hasta abril de 2004

**Anexo 2:** Muestra las notificaciones procesadas en el sistema de Farmacovigilancia nacional en el período 2005-2009.

Año	No. reportes	Tasa reporte anual
2005	7025	624 x 1 000 000
2006	8261	735 x 1 000 000
2007	6928	615 x 1 000 000
2008	7203	640 x 1 000 000
2009	8053	716 x 1 000 000

**Anexo 3:** Muestra la distribución de efectos adversos de baja frecuencia de aparición según severidad en Cuba en el periodo 2004-2008.

Frecuencia RAM	Leve		Moderada		Grave		Mortal		Total	
	No.	%	No.	%	No.	%	No.	%	No	%
Ocasional	3204	26.3	3878	31.8	162	1.3	23	0.2	7267	59.6
Rara	1280	10.5	1929	15.8	128	1	29	0.2	3366	27.6
No descrita	698	5.7	787	6.4	59	0.5	13	0.1	1557	12.8
Total	5182	42.5	6594	54.1	349	2.8	65	0.5	12190	100

**Anexo 4:** Muestra un estudio de las RAM por distribución de notificaciones de estas según su severidad.

Severidad	2005		2006		2007		2008		2009	
	No	%	No	%	No	%	No	%	No	%
Leves	4218	55.0	3550	51.2	4547	60.0	3298	45.8	3343	41.5
Moderadas	2699	42.5	3262	47.1	3510	38.4	3784	52.5	4586	56.9
Graves	85	2.2	95	1.4	185	1.2	105	1.5	86	1.1
Mortales	23	0.2	20	0.3	19	0.3	16	0.2	38	0.5

**Anexo 5:** Muestra las reacciones adversas evitables en Hospitales seleccionados de Ciudad Habana durante el 2006.

Causas	Nº	%	Fármaco
Automedicación	4	14,8	aspirina, ibuprofeno, piroxicam, naproxeno
Indicación inadecuada	6	22,2	ibuprofeno, penicilina cristalina, amoxicodex, tioridazina
Dosis inadecuada	2	7,4	penicilina cristalina, piroxicam
Error en la pauta del tratamiento	4	14,8	ibuprofeno, piroxicam, dipirona, indometacina
Asociaciones inadecuadas	10	37,0	aspirina + indometacina + naproxeno ibuprofeno + aspirina paracetamol + ibuprofeno paracetamol + ibuprofeno + amicodex ibuprofeno + dipirona + tioridazina + trifluoperazina

**Anexo 6:** Refleja los costos de los servicios en los que ingresaron pacientes con RAM en el Hospital Joaquín Albarrán de la Ciudad Habana en el año 2006.

Partidas	Cuerpo Guardia	Cirugía	Cardiología	Terapias	Dermatología	Costo Total	%
Salario	\$217.263	\$853.763	\$69.488	\$489.301	\$41.903	\$1.671.718	37,6
Gasto Material	\$283.537	\$499.886	\$47.610	\$351.886	\$91.057	\$1.273.976	28,6
Otros	\$187.224	\$226.630	\$30.859	\$156.330	\$26.426	\$627.469	14,1
Medicamentos	\$192.416	\$289.667	\$25.495	\$301.806	\$ 65.359	\$874.743	19,7
TOTAL	\$880.440	\$1.869.946	\$173.452	\$1.299.323	\$ 224.745	\$4.447.906	100

**Anexo 7:** Muestra el número de casos y las estimaciones nacionales de las visitas a urgencias y hospitalizaciones de emergencia por RAM en adultos mayores, según las características del paciente y la sentencia, 2007-2009.

Características	Hospitalizaciones			Visitas a urgencias sin hospitalización		
	No de casos (N= 5077)	Estimado Nacional Anual (N = 99,628)		No de casos (N= 7,589)	Estimado Nacional Anual (N = 166,174)	
		no	%(95 % C)		no	%(95 % C)
<b>Años</b>						
65–69 años	801	14,179	14.2 (12.0–16.5)	1669	36,380	21.9 (19.7–24.1)
70–74 años	924	18,257	18.3 (16.6–20.1)	1546	32,575	19.6 (18.4–20.8)
75–79 años	1001	19,248	19.3 (18.2–20.5)	1628	35,702	21.5 (20.1–22.9)
80–84 años	1110	22,619	22.7 (20.9–24.5)	1366	31,266	18.8 (17.2–20.4)
≥85 años	1241	25,326	25.4 (23.0–27.9)	1380	30,251	18.2 (16.6–19.8)
<b>Sexo</b>						
Femenino	2969	59,278	59.5 (57.4–61.6)	4511	99,495	59.9 (57.7–62.0)
Masculino	2106	40,302	40.5 (38.4–42.5)	3076	66,604	40.1 (37.9–42.3)
<b>Tipos de Reacciones Adversas</b>						
Sobredosis no intencional	3375	65,450	65.7 (60.1–71.3)	3608	75,982	45.7 (40.7–50.7)
Efectos adversos	1390	27,613	27.7 (22.3–33.1)	2313	50,240	30.2 (26.0–34.4)
Reacción alérgica	267	5,617	5.6 (3.4–7.9)	1388	33,838	20.4 (17.0–23.7)
Otros	45	948	1.0 (0.7–1.2)	280	280 6,115	3.7 (2.9–4.4)
<b>No. de medicaciones implicadas</b>						
1	4204	82,050	82.4 (78.5–86.3)	6471	141,939	85.4 (82.5–88.3)
2	873	17,578	17.6 (13.7–21.5)	1118	24,235	14.6 (11.7–17.5)
<b>No. de medicaciones concomitantes</b>						
Ninguna	773	18,324	18.4 (12.0–24.8)	2103	56,082	33.7 (22.8–44.7)
1 - 4	1459	26,731	26.8 (21.9–31.8)	2156	43,819	26.4 (21.2–31.5)

5 – 9	2115	40,443	40.6 (35.0–46.2)	2554	50,420	30.3 (24.0–36.7)
>=10	730	14,130	14.2 (9.2–19.1)	776	15,853	9.5 (6.8–12.3)

**Anexo 8:** Muestra la diferencia significativa que existe en los tiempos de cálculo del BT y el LEX a medida que aumenta la dimensión de las matrices.

Orden de MB	CC	CT	BT	LEX	# TT
22x15	14832 mseg.	571 mseg.	30 mseg.	10 mseg.	165
24x16	39697 mseg.	2824 mseg.	70 mseg.	30 mseg.	230
29x17	3 min.	4457 mseg.	200 mseg.	50 mseg.	430
54x20	1 h.	4 min.	2424 mseg.	481 mseg.	2294
59x25		2 h.	47999 mseg.	2103 mseg.	11746
97x25			60033 mseg.	8172 mseg.	19973
58x30			10 min.	30 seg.	34900
65x35			> 90 in.	1 min.	112219
70x40			> 150 min	2 min.	446986
100x60			> 12 h.	2 h.	16567230