



**Universidad Estatal a Distancia**  
**Maestría en Administración de Negocios**  
**Mención en Gerencia de Proyectos**

**Trabajo Final de Graduación**

**Modelo de redes neuronales optimizadas con algoritmos  
genéticos.**

**Una aplicación para proyecciones sobre pacientes con cáncer  
hospitalizados en la C.C.S.S.**

**Estudiantes:**

Andrey Alonso Mora Bermúdez

Juan Antonio Sancho Chacón

**Directora del proyecto:** MSc. Ana Lorena Solís Guevara

**Lector:** Dr. Róger Méndez Benavides

**Enero, 2012**

## **Dedicatorias**

A Paola que ha sido una bendición en mi vida, porque con su apoyo incondicional, entrega y amor ha sido mi motivación e inspiración en esta nueva etapa de mi formación profesional. A mis padres, a quienes les debo todo lo que soy y lo que tengo; todos mis logros han sido y serán por y para ellos. Finalmente a mis hermanos y a mis hermanas, porque toda mi vida he podido contar con ellos y me enorgullece tenerlos como familia. Los quiero.

ANDREY

A Dios, quien me ha dado la vida y la oportunidad de culminar de manera exitosa esta nueva etapa. A mi familia quienes a pesar de la distancia y las dificultades siempre me ha apoyado y de manera especial a Milena, una mujer maravillosa que Dios ha enviado a mi vida, quien ha sabido comprenderme, animarme y apoyarme en este proceso y sobre todo que ha iluminado mi vida con su ternura y su amor sincero.

JUAN ANTONIO

## **Agradecimientos**

Queremos agradecer a la Dra. Ana Lorena Solís Guevara, nuestra Directora de Tesis, porque con su guía y orientación hemos logrado concluir exitosamente nuestro proyecto de graduación.

Agradecemos a la Caja Costarricense de Seguro Social, en especial al Área de Estadísticas en Salud, por facilitarnos los datos y herramientas que nos permitieron desarrollar esta investigación, que pretende contribuir con la mejora de la calidad en la prestación de servicios de salud.

Un agradecimiento especial a la Dra. Julia Lee Vargas, ya que en su labor tanto de profesora como de Directora de la Maestría, nos orientó responsablemente y nos dedicó su tiempo y esfuerzo para que pudiéramos alcanzar nuestra meta.

También agradecemos a la Universidad Estatal a Distancia y a los profesores de la Maestría en Administración de Negocios, por su aporte en nuestra formación profesional y compromiso con ofrecer una educación de calidad. Mención especial merece nuestro Lector de Tesis, el Dr. Róger Méndez Benavides, porque nos apoyó e impulsó en la idea que dio origen a este proyecto de graduación.

Finalmente, queremos dar gracias a Dios, por darnos el don de la vida, por permitirnos ser lo que somos y por regalarnos una bendición más como lo es esta maestría.

## Resumen Ejecutivo

En el campo de la salud, adquiere fundamental importancia disponer de proyecciones efectivas de la demanda de los servicios de salud, con el fin de realizar una adecuada planificación, garantizando la disposición oportuna de dichos servicios y una utilización eficiente de los recursos.

El objetivo de la presente investigación es ampliar el conocimiento teórico en relación con la utilización de las redes neuronales y los algoritmos genéticos como modelos de proyección y optimización, con el fin de evaluar su aplicación práctica en la estimación del número de pacientes con cáncer atendidos en hospitales de la CCSS y la estancia hospitalaria esperada para este tipo de paciente.

Para alcanzar este objetivo se desarrollaron dos redes neuronales, una para la proyección de casos nuevos de pacientes con cáncer hospitalizados por mes (red de casos nuevos) y otro para determinar la estancia hospitalaria (red de estancia).

Mediante el análisis de las variables relacionadas con estos dos fenómenos de estudio, con base en las recomendaciones teóricas para la construcción de redes neuronales y utilizando dos programas estadísticos (SPSS y Statistica) se determinó que la estructuración de la red de casos nuevos debía ser 12:7:1, es decir una capa de entrada con 12 variables o neuronas, una capa oculta con 7 neuronas y una capa de salida con una neurona. Por su parte la estructuración de la red de estancia recomendada es 8:7:1.

Para la red de casos nuevos, las variables de entrada seleccionadas fueron: tasa de mortalidad, índice de desarrollo humano, inversión social en salud, cobertura régimen de enfermedad y maternidad, tasa de desempleo, nivel de pobreza, mes, promedio de casos nuevos de pacientes con cáncer hospitalizados en los 3 últimos años, promedio de los 2 últimos años, promedio del último año, promedio de los últimos 6 meses y promedio de los últimos 3 meses.

En el caso de la red de estancia se seleccionaron las siguientes 8 variables de entrada: edad del paciente, sexo, residencia, centro de salud, área de ingreso, servicio de ingreso, ingresos previos y diagnóstico principal.

Posteriormente se aplicaron los principios teóricos y las formulas matemáticas de redes neuronales para la construcción y entrenamiento de ambos modelos. La metodología de aprendizaje utilizada en ambas redes es la propagación del error hacia atrás (*back-propagation*).

La red de casos nuevos, se entrenó con un total de 120 casos (71% de los disponibles). Para analizar la efectividad del entrenamiento se utilizaron 36 datos (21%) y el restante 8% se utilizó para comparar las proyecciones. El resultado del porcentaje de error medio absoluto (MAPE, por sus siglas en inglés) durante el entrenamiento fue de 11.76% y después del entrenamiento el error fue de 7.53%

En el caso de la red de estancia, para el entrenamiento se utilizaron los últimos 1000 casos disponibles. De estos, 700 (70%) se utilizaron para el entrenamiento, 150 (15%) para la evaluación y 150 (15%) para comparar las proyecciones. El resultado del error total durante el entrenamiento fue de 84.13% y después del entrenamiento el error fue de 47.37%.

Una vez construidos los modelos de redes neuronales, el siguiente paso fue la optimización mediante algoritmos genéticos. El algoritmo definido tanto para la red de casos nuevos como para la de estancia hospitalaria, es un algoritmo de minimización del error de pronóstico o estimación.

Después de la optimización el resultado del error para la red de casos nuevos pasó de 7.53% a 5.07%. En el caso de la red de estancia el resultado del error pasó de 47.37% a 40,79%. En ambos casos se logra apreciar la efectividad del proceso de optimización, ya que, se lograron reducciones importantes en los porcentajes de error de ambas redes.

Para analizar la efectividad del modelo de red neuronal de casos nuevos, se generaron proyecciones para los siguientes 12 períodos y se compararon con las proyecciones obtenidas mediante 4 modelos tradicionales de pronósticos: suavización exponencial simple, suavización exponencial doble, “*winters*” aditivo y “*winters*” multiplicativo. Con esto se determinó que el modelo de redes neuronales genera los mejores resultados en términos del error de pronóstico (5.07%). El más cercano de los otros métodos fue el Winters Aditivo con un error de 6.24%.

Finalmente, se evaluó el horizonte de pronóstico de la red de casos nuevos analizando tres escenarios: un año, dos años, y tres años. Los resultados de la red neuronal son más efectivos que los obtenidos con los modelos tradicionales en los tres años analizados, donde por ejemplo en el primer año el MAPE de la red es de 5.63%, muy por debajo de 6.9% del SES que es el mejor de los tradicionales, y para el tercer año el error de la red neuronal es de 5.22%, siendo menor que el obtenido por los modelos tradicionales, cuyo mejor resultado fue un 7,2%.

Como conclusión principal de la investigación se determina que la utilización de un modelo de redes neuronales, optimizadas con algoritmos genéticos, para proyectar la cantidad de casos nuevos de pacientes con cáncer hospitalizados en la CCSS, genera mejores resultados en términos del menor error de pronóstico, que los obtenidos utilizando algunas de las técnicas tradicionales de pronósticos.

El modelo desarrollado presenta mejores resultados que los métodos tradicionales considerando un horizonte de pronósticos de uno, dos o tres años. Después de este período se recomienda hacer una actualización de datos de entrada de la red, así como un nuevo entrenamiento y optimización, ya que, los cambios en variables macroeconómicas y sociales incorporadas en el modelo se manifiestan con un rezago de tres años.

## Contenido

INTRODUCCIÓN .....	15
CAPÍTULO I. ANTECEDENTES Y OBJETIVOS DE INVESTIGACIÓN .....	18
1.1 Antecedentes del problema de investigación .....	18
1.2 Justificación del problema de investigación .....	22
1.3 Formulación del problema de investigación .....	24
1.4 Delimitaciones del estudio .....	24
1.5 Objetivos .....	25
1.5.1 Objetivo general .....	25
1.5.2 Objetivos específicos .....	25
1.6 Hipótesis .....	26
CAPÍTULO 2. MARCO REFERENCIAL.....	27
2.1 Marco contextual.....	27
2.1.1 Pronósticos .....	27
2.1.2 Redes Neuronales Artificiales.....	28
2.1.3 Algoritmos genéticos (AG).....	30
2.1.4 Investigaciones conjuntas .....	31
2.2 Marco Conceptual.....	32
2.2.1 Pronósticos .....	32
2.2.2 Redes Neuronales Artificiales.....	43
2.2.3 Algoritmos genéticos (AG).....	60
2.3 Marco Teórico .....	66
CAPÍTULO III. MARCO METODOLÓGICO.....	69
3.1 Tipo de estudio .....	69
3.2 Sujetos y fuentes.....	72
3.2.1 Sujetos .....	72
3.2.2 Fuentes .....	72

3.3	Tiempo y espacio del proyecto .....	73
3.4	Variables .....	74
3.5	Tipo de muestreo .....	80
3.6	Descripción de los instrumentos .....	81
3.6.1	Entrevista no estructurada .....	81
3.6.2	Listado o base de datos históricos .....	81
3.6.3	Aplicación automatizada del modelo .....	82
3.7	Tratamiento de la información .....	82
CAPÍTULO IV. DESARROLLO DE LA INVESTIGACIÓN .....		84
4.1	Descripción del fenómeno en estudio .....	84
4.1.1	Evolución número de casos nuevos de pacientes hospitalizados con cáncer	85
4.1.2	Sexo .....	86
4.1.3	País de nacimiento.....	86
4.1.4	Edad .....	87
4.1.5	Residencia.....	88
4.1.6	Centros de salud de ingreso .....	89
4.1.7	Área de ingreso.....	91
4.1.8	Servicio de ingreso.....	92
4.1.9	Especialidad de ingreso .....	92
4.1.10	Diagnóstico principal .....	93
4.1.11	Estancia preoperatoria .....	95
4.1.12	Procedimientos desarrollados.....	96
4.1.13	Condición de salida.....	97
4.1.14	Causa básica de muerte.....	98
4.1.15	Estancia total .....	99
4.1.16	Variables macroeconómicas y sociales que afectan el fenómeno.....	100
4.2	Selección de las variables de entrada .....	110
4.2.1	Variables de entrada para la red de casos nuevos.....	110

4.2.2	Variables de entrada para la red de estancia.....	116
4.3	Diseño del modelo de redes neuronales.....	119
4.3.1	Tipo de red.....	119
4.3.2	Estructuración de la red.....	119
4.3.3	Funcionamiento de la red.....	126
4.3.4	Entrenamiento de la red.....	129
4.3.5	Evaluación inicial de la red.....	134
4.4	Optimización de la red mediante algoritmos genéticos (AG).....	136
4.4.1	Tamaño de población.....	137
4.4.2	Población inicial.....	137
4.4.3	Evaluación de individuos.....	138
4.4.4	Selección.....	139
4.4.5	Recombinación o cruce.....	140
4.4.6	Mutación.....	142
4.4.7	Evaluación de nueva generación.....	142
4.4.8	Evaluación de la red neuronal optimizada.....	143
4.4.9	Generación de pronósticos mediante el modelo de redes neuronales.....	145
4.5	Comparación de resultados de la red neuronal con los generados por un programa informático especializado en pronósticos.....	147
4.5.1	Evaluación del horizonte del pronóstico.....	151
4.6	Diseño de una herramienta en Excel para utilizar la red neuronal en proyecciones de pacientes con cáncer.....	153
CAPÍTULO V. APORTES TEÓRICOS Y METODOLÓGICOS DE LA INVESTIGACIÓN.....		154
CONCLUSIONES.....		159
RECOMENDACIONES.....		162
REFERENCIAS BIBLIOGRÁFICAS.....		163
ANEXOS.....		168

## Índice de cuadros

Cuadro 1: Algunas aplicaciones de las RNA.....	45
Cuadro 2: Variables de la investigación.....	75
Cuadro 3: Variables internas del modelo.....	79
Cuadro 4: Principales cantones de residencia de los pacientes nuevos con cáncer hospitalizados en la CCSS en el año 2010.....	89
Cuadro 5: Distribución del número de casos de pacientes con cáncer atendidos en la CCSS, según servicio de ingreso. ....	92
Cuadro 6: Distribución del número de casos de pacientes con cáncer atendidos en la CCSS, según condición de salida.....	97
Cuadro 7: Distribución del número de casos de pacientes con cáncer hospitalizados en la CCSS, según estancia total en el período 2000-2010. ....	99
Cuadro 8: Determinación de período de incidencia de variables macro, con Statistica. ..	112
Cuadro 9: Análisis de correlaciones entre las variables de entrada para la Red de casos nuevos, con SPSS. ....	114
Cuadro 10: Priorización de variables de entrada Red de casos nuevos, con SPSS. ....	115
Cuadro 11: Análisis de correlaciones entre las variables de entrada para la Red de estancia, con SPSS.....	118
Cuadro 12: Ejemplo de neuronas de entrada en la Red de casos nuevos .....	126
Cuadro 13: Pesos y entradas netas en la Red de casos nuevos .....	127
Cuadro 14: Cálculo de la entrada neta para la neurona oculta $O_{h1}$ en la Red de casos nuevos .....	128
Cuadro 15: Cálculo de una salida final en la Red de casos nuevos .....	129
Cuadro 16: Cálculo del error para las neuronas ocultas Red de casos nuevos.....	131
Cuadro 17: Cálculo del error para las neuronas ocultas Red de casos nuevos.....	132
Cuadro 18: Cálculo del error para las neuronas de entrada en la Red de casos nuevos ...	133

Cuadro 19: Cálculo del $\Delta w_{ih}$ para la neurona de entrada 2 (Prom 3a) en la Red de casos nuevos .....	133
Cuadro 20: Evaluación inicial de las redes de casos nuevos y de estancia hospitalaria ....	135
Cuadro 21: Definición de la población inicial .....	138
Cuadro 22: Evaluación de la población inicial .....	139
Cuadro 23: Selección de los individuos .....	140
Cuadro 24: Evaluación de los nuevos individuos .....	141
Cuadro 25: Detalle de la mutación de los individuos .....	142
Cuadro 26: Evaluación de la nueva generación .....	143
Cuadro 27: Evaluación inicial de las redes de casos nuevos y de estancia hospitalaria ....	144
Cuadro 28: Proyecciones de casos nuevos de pacientes hospitalizados con cáncer utilizando para 12 períodos de análisis (año 2011).....	145
Cuadro 29: Comparación de resultados de pronósticos generados por los modelos tradicionales y el modelo de redes neuronales .....	148
Cuadro 30: Comparación de errores de pronósticos generados por los modelos tradicionales y el modelo de redes neuronales .....	150
Cuadro 31: Evaluación del horizonte de pronósticos del modelo de redes neuronales en comparación con los modelos tradicionales .....	151

## Índice de gráficos

Gráfico 1 : Evolución del número de casos nuevos de pacientes hospitalizados con cáncer atendidos en la CCSS entre los años 2000 y 2010.....	85
Gráfico 2 : Distribución porcentual de la cantidad de casos nuevos de pacientes con cáncer hospitalizados en la CCSS, según sexo para el período 2000-2010.....	86
Gráfico 3 : Distribución del número de casos nuevos de pacientes con cáncer hospitalizados en la CCSS, según edad para el período 2000-2010.....	87
Gráfico 4 : Distribución porcentual de la cantidad de casos nuevos de pacientes con cáncer hospitalizados en la CCSS, según provincia de residencia para el período 2000-2010. ....	88
Gráfico 5 : Centros de salud de la CCSS en los que son hospitalizados la mayor cantidad de pacientes nuevos con cáncer, en el período 2000-2010.....	90
Gráfico 6 : Distribución del número de casos de pacientes nuevos con cáncer hospitalizados en la CCSS, según área de ingreso para el período 2000-2010.....	91
Gráfico 7 : Principales especialidades de ingreso de los pacientes con cáncer hospitalizados en la CCSS en el período 2005-2010.....	93
Gráfico 8 : Diagnósticos principales que presentaron la mayor cantidad de pacientes con cáncer hospitalizados en la CCSS en el período 2000-2010.....	94
Gráfico 9 : Distribución del número de casos nuevos de pacientes con cáncer hospitalizados que requirieron cirugía, según la estancia preoperatoria para el período 2000-2010.....	95
Gráfico 10 : Principales procedimientos realizados a los pacientes con cáncer hospitalizados en la CCSS en el período 2000-2010.....	96
Gráfico 11: Principales causas de muerte de los pacientes con cáncer hospitalizados en la CCSS en el período 2000-2010. ....	98
Gráfico 12: Esperanza de vida al nacer en Costa Rica para el período 2000-2010 (en años). .....	100

Gráfico 13: Tasa general de mortalidad en Costa Rica para el período 2000-2010 (en número de defunciones por cada mil habitantes). .....	101
Gráfico 14: Población total de Costa Rica para el período 2000-2010 (en número de personas). .....	102
Gráfico 15: Índice de desarrollo humano en Costa Rica para el período 2000-2010. ....	103
Gráfico 16: Inversión social en salud en Costa Rica para el período 2000-2010 (porcentaje del PIB).....	104
Gráfico 17: Tasa de escolaridad y nivel educativo en Costa Rica para el período 2000-2010 (porcentaje) .....	105
Gráfico 18: Cobertura Régimen de Enfermedad y Maternidad en Costa Rica para el período 2000-2010 (porcentaje de la población nacional). .....	106
Gráfico 19: Desigualdad en la distribución del ingreso en Costa Rica para el período 2000-2010 (coeficiente de Gini). .....	107
Gráfico 20: Tasa de desempleo en Costa Rica para el período 2000-2010.....	108
Gráfico 21: Nivel de pobreza en Costa Rica para el período 2000-2010.....	109
Gráfico 22: Proyecciones de casos nuevos de pacientes hospitalizados con cáncer para 12 períodos de análisis (año 2011). .....	145
Gráfico 23 : Comparación de resultados de pronósticos generados por los modelos tradicionales y el modelo de redes neuronales .....	149

## Índice de figuras

Figura 1: Partes de una neurona biológica.....	47
Figura 2: Arquitectura de una red neuronal artificial.....	48
Figura 3: Funciones de transferencia utilizadas en los modelos de Redes Neuronales Artificiales .....	54
Figura 4: Resultados del modelado de la red de casos nuevos con Statistica .....	121
Figura 5: Resultados del modelado de la red de casos nuevos con SPSS .....	122
Figura 6: Resultados del modelado de la red de estancia con Statistica .....	122
Figura 7: Resultados del modelado de la red de estancia con SPSS .....	123
Figura 8: Estructuración de la red neuronal de casos nuevos de pacientes con cáncer....	124
Figura 9: Estructuración de la red neuronal de estancia de los pacientes con cáncer. ....	125

## Listado de abreviaturas

**AG:** Algoritmo Genético.

**CCSS:** Caja Costarricense de Seguro Social.

**CIE-10:** Clasificación Internacional de Enfermedades.

**CREM:** Cobertura del Régimen de Enfermedad y Maternidad.

**Fa:** Frecuencia de aprendizaje.

**IDH:** Índice de Desarrollo Humano.

**INEC:** Instituto Nacional de Estadística y Censos.

**ISS:** Inversión Social en Salud.

**MAPE:** Mean Absolute Porcentual Error (Error Porcentual Medio Absoluto).

**MAD:** Mean Absolute Desviation (Desviación Media Absoluta).

**MIDEPLAN:** Ministerio de Planificación.

**MSE:** Mean Square Error (Error Medio Cuadrado).

**NP:** Nivel de pobreza.

**PIB:** Producto Interno Bruto.

**PMP:** Promedio Móvil Ponderado.

**PMS:** Promedio Móvil Simple.

**PNUD:** Programa de Naciones Unidas para el Desarrollo.

**PP:** Promedio Ponderado.

**Prom:** Promedio.

**RN:** Red Neuronal.

**RNA:** Red Neuronal Artificial.

**SED:** Suavización Exponencial Doble.

**SES:** Suavización Exponencial Simple.

**Sf:** Sin fecha de publicación.

**Snp:** Sin número de página.

**SPSS:** Statistical Package for the Social Sciences.

**TD:** Tasa de desempleo.

**TM:** Tasa de mortalidad.

**VBA:** Visual Basic Applications.

**Vol:** Volumen.

**WA:** Winters Aditivo.

**WM:** Winters Multiplicativo.

## Glosario de términos

**Algoritmo:** Conjunto de procedimientos o pasos ordenados que permiten convertir ciertos valores de entrada en salidas que funcionan como solución a diversos problemas o al menos forman parte de la solución, especialmente aplicado a problemas matemáticos.

**Correlación:** Es la magnitud y dirección de una relación lineal entre dos variables.

**Back-propagation:** Método de propagación del error hacia atrás utilizado como algoritmo de entrenamiento de una red neuronal.

**Cáncer:** Conjunto de enfermedades en las cuales las células anormales se dividen sin control y pueden invadir otros tejidos.

**Capa:** Conjunto de neuronas o nodos de una red neuronal que se encuentra ubicados en un mismo nivel de la red.

**Coefficiente de concentración de Gini:** Es una medida de la desigualdad en la distribución del ingreso, donde el valor 0 corresponde a una igualdad total (todos tienen lo mismo) y el valor 1 la perfecta desigualdad (uno tiene todo y los demás nada).

**Horizonte de pronóstico:** Corresponde al periodo o espacio de tiempo para el que se generan proyecciones con base en un modelo de pronósticos determinado.

**Iteración:** Es un conjunto de actividades o pasos de un proceso que se repiten en el tiempo, utilizado especialmente para la solución de problemas del tipo matemáticos o lógicos.

**Minitab:** Programa informático para el análisis estadístico.

**Modelo:** Organización e interrelación de un conjunto de elementos que intervienen en un fenómeno de estudio determinado para generar un resultado.

**Neurona:** Unidad básica de constitución de una red neuronal que permite la comunicación entre los elementos de entrada y los elementos de salida de la red. También se hace referencia a las neuronas biológicas que constituyen el sistema nervioso.

**Nodo:** Es el espacio donde confluyen las conexiones de una red neuronal, coincide con las neuronas de las capas intermedias y de salida de la red.

**Oncología:** Especialidad médica que estudia y trata las enfermedades del cáncer y los procesos similares como neoplasias y tumores.

**Optimización:** Proceso de buscar los mínimos y máximos de diferentes funciones matemáticas.

**Sobrevida:** Tiempo de vida de una persona más allá de lo esperado.

**Statistica:** Programa informático para el análisis estadístico.

## INTRODUCCIÓN

Los pronósticos son estimaciones anticipadas del valor de una o más variables. Constituyen una herramienta fundamental que disponen las organizaciones para realizar una adecuada planificación estratégica y tomar mejores decisiones, sobre todo en temas que requieren una atención especial y cuyo comportamiento depende de muchas otras condiciones, incluidas sociales y económicas.

Con el paso del tiempo se han ido desarrollando diversas técnicas de pronósticos, unas cualitativas y otras cuantitativas, que toman en consideración distintas variables relacionadas con el comportamiento de los datos. Asimismo, con el avance a nivel tecnológico, se ha diseñado aplicaciones que permiten realizar pronósticos con el fin de facilitarle esta tarea a las organizaciones. Sin embargo, aunque muchas de ellas utilizan estos programas computacionales, el grado de efectividad de los mismos es variable, ya que, los supuestos bajo los cuales trabajan no siempre se adecuan a los datos que se están utilizando.

En los últimos años, los esfuerzos por mejorar la calidad de los pronósticos han dado paso a nuevas técnicas basadas en la lógica del funcionamiento del cerebro humano y de la evolución de los seres vivos, estas son respectivamente las redes neuronales y los algoritmos genéticos.

Las redes neuronales artificiales son modelos matemáticos que tratan de hacer una pequeña emulación del funcionamiento del cerebro humano y permiten hacer pronósticos donde hay cierto comportamiento no lineal de los datos. Estas redes se “aprenden” la dinámica del fenómeno en estudio y tienen en cuenta el tiempo como una variable que afecta el fenómeno.

Por su parte los algoritmos genéticos son métodos adaptativos que basan su funcionamiento en los procesos genéticos de evolución de los seres vivos, donde los

individuos mejor adaptados son los que logran sobrevivir y reproducirse, haciendo que sus genes perduren, esta dinámica es utilizada por estos algoritmos para ir creando soluciones que vayan optimizando los valores de ciertas variables, con los que se puedan obtener mejores resultados para el problema que se está analizando.

Se han realizado algunos estudios que indican que el aprendizaje de las redes neuronales puede optimizarse utilizando algoritmos genéticos, obteniendo de esta forma resultados más robustos y cercanos a la realidad.

La finalidad de este trabajo es ampliar el conocimiento teórico en relación con la utilización de las redes neuronales y los algoritmos genéticos en el desarrollo de modelos de proyección y optimización, con el fin de evaluar su aplicación práctica en la estimación del número de pacientes con cáncer atendidos en hospitales de la CCSS y la estancia hospitalaria esperada para este tipo de paciente.

La estructuración de este trabajo es la siguiente:

- Capítulo 1. Definición del problema y objetivos: comprende la definición, antecedentes y justificación del tema. Además en este capítulo se definen los objetivos del proyecto, se plantea el problema y se establece la hipótesis de investigación.
- Capítulo 2. Marco referencial: en este capítulo se abordan detalladamente, según diferentes actores, la temática de la investigación. Para el caso del presente proyecto se abarcan tres temáticas prioritarias: pronósticos, redes neuronales y algoritmos genéticos.
- Capítulo 3. Marco metodológico: se define el tipo de investigación que se desarrolló, así como la metodología que se utilizó para alcanzar los diferentes objetivos del trabajo.

- Capítulo 4. Desarrollo de la investigación: comprende el diseño del modelo de redes neuronales a partir del análisis de las variables que afectan el fenómeno de estudio. Además se realiza la optimización del modelo mediante algoritmos genéticos, y por último se comparan los resultados que genera el modelo de redes neuronales con los obtenidos mediante otras técnicas tradicionales de pronósticos, con el fin de determinar cuál es método es más efectivo.
- Capítulo 5. Aportes teóricos y metodológicos: como resultado del desarrollo de la investigación se realizan recomendaciones a nivel del procedimiento a seguir cuando se utilizan las técnicas de redes neuronales y algoritmos genéticos de forma combinada.

## **CAPÍTULO I. ANTECEDENTES Y OBJETIVOS DE INVESTIGACIÓN**

El objetivo del primer capítulo es establecer los antecedentes y objetivos de la investigación. El mismo está estructurado en seis apartados: antecedentes, justificación, formulación del problema, delimitaciones del estudio, objetivos e hipótesis.

### **1.1 Antecedentes del problema de investigación**

La necesidad del ser humano de pronosticar eventos es muy antigua, y se da debido a varios factores como estar en un entorno altamente incierto, en el que la intuición no da los mejores resultados, y en el que se ha requerido anticipar o prever el comportamiento futuro de ciertos eventos con el fin de planificar y tomar decisiones efectivas.

En el mundo complejo como el que existe hoy y bajo influencias de crisis financieras internacionales, las organizaciones se han visto obligadas a buscar mayor eficiencia en sus procesos de trabajo. Un pronóstico acertado y una adecuada planeación de la demanda de bienes o servicios, en cualquiera de sus fases, permitirían controlar los costos totales de operación mediante la distribución de los esfuerzos de acuerdo a las necesidades encontradas. Esto se hace aún más relevante cuando las organizaciones atraviesan situaciones coyunturales como una crisis financiera interna, tal es el caso de la Caja Costarricense de Seguro Social.

Cuando no se elabora una proyección de la demanda o se hace de forma incorrecta, la planeación se ve afectada y los esfuerzos son orientados hacia aspectos que no son los de más impacto, desaprovechando recursos que normalmente hacen falta en otras áreas más relevantes. En el caso de la salud, si la demanda de servicios en algún tema especial se subestima, las consecuencias se reflejan en la disponibilidad de servicios para la población, afectando de forma importante el acceso y la oportunidad.

La estimación del comportamiento futuro de algunas variables puede realizarse utilizando diversas técnicas de pronóstico, algunas técnicas son tradicionales y otras son de reciente surgimiento.

De acuerdo con Nojek, Britos, Rossi y García Martínez (2003) una forma de clasificar las técnicas tradicionales de proyección consiste en hacerlo en función de su carácter, esto es, aplicando métodos cualitativos, modelos causales y modelos de series de tiempo.

- Los métodos de carácter cualitativo se basan principalmente en opiniones de expertos. Su uso es frecuente cuando el tiempo para elaborar el pronóstico es escaso, cuando no se dispone de todos los antecedentes mínimos necesarios o cuando los datos disponibles no son confiables para predecir el comportamiento futuro.
- Los modelos de pronóstico causales parten del supuesto de que el grado de influencia de las variables que afectan al comportamiento del mercado permanece estable, para luego, construir un modelo que relacione ese comportamiento con las variables que se estima que son las causantes de los cambios que se observan en el mercado.
- Los modelos de series de tiempo se refieren a la medición de valores de una variable en el tiempo, a intervalos espaciados uniformemente. El objetivo de la identificación de la información histórica es determinar un patrón básico en su comportamiento, que posibilite la proyección futura de la variable deseada.

Dentro de los métodos tradicionales de pronóstico, es decir aquellos más frecuentemente utilizados, se pueden mencionar a nivel general: juicio de experto, análisis de series de tiempo, análisis de regresión y pruebas de mercado. Además, hay métodos específicos como la suavización exponencial simple y doble, Winters aditivo y multiplicativo, Box-Jenkins, entre otros.

Cada una de las técnicas de proyección tiene una aplicación de carácter especial que hace de su selección un problema de decisión influido por diversos factores, como por ejemplo, la validez y disponibilidad de los datos históricos, la precisión deseada del pronóstico, el costo del procedimiento, los beneficios del resultado, los períodos futuros que se desee pronosticar y el tiempo disponible para hacer el estudio, entre otros (Sapag Chain, 2000).

Muchas de las técnicas de pronóstico que se utilizan actualmente se desarrollaron en el siglo XIX. Antes del surgimiento de las computadoras, el cálculo de pronósticos era una labor compleja, y aunque existían los mecanismos para hacerlo sin computadora, muchas organizaciones preferían utilizar la intuición de su personal como técnica de proyección.

Junto con la aparición de las computadoras se inició el desarrollo de técnicas de pronóstico más complejas, ocasionando que en los últimos años los pronósticos recibieron más atención, incrementándose la confianza en las técnicas que abarcan una compleja manipulación de datos.

Así surgieron diferentes programas computacionales para pronósticos, cuyo objetivo fue facilitar esta labor a las organizaciones, mediante una interfaz amigable que les permitiera introducir los datos y obtener las predicciones fácilmente. Se han desarrollado distintas *aplicaciones* o programas computacionales con base en las diferentes técnicas de pronóstico, la diferencia radica en la calidad de proyecciones obtenidas, medida a través del menor error de pronóstico, entendiendo este como la diferencia esperada entre el valor proyectado y el real.

Algunos programas computacionales funcionan mejor que otros dependiendo de la naturaleza de las operaciones de la organización y del comportamiento esperado de sus datos; sin embargo, este comportamiento de los datos en ocasiones varía, por lo que una organización podría estar utilizando actualmente un modelo o aplicación de pronósticos que en el mediano plazo podría no ser funcional.

La tecnología y el conocimiento no han dejado de avanzar, lo que ha permitido que recientemente se estén desarrollando nuevos modelos y técnicas, para resolver problemas de las organizaciones que se tornaban complejos utilizando técnicas tradicionales.

Valenzuela (2008) se refiere a algunas de estas nuevas técnicas como sistemas basados en la inteligencia, los cuales se fundamentan en la forma en que funciona el cerebro humano o en la lógica de la evolución. Entre estas herramientas se incluyen: computación evolutiva, algoritmos genéticos y las redes neuronales; los cuales han sido utilizados en diferentes aplicaciones incluyendo el pronóstico de diversas variables que afectan el accionar de las organizaciones.

Dentro de los métodos no tradicionales de pronósticos se encuentran las redes neuronales. De acuerdo con Toro, Mejía y Salazar (2004) las redes neuronales artificiales son modelos matemáticos que tratan de hacer una pequeña emulación del funcionamiento del cerebro humano. Son metodologías novedosas que permiten hacer pronósticos donde hay cierto comportamiento no lineal. Las redes neuronales se “aprenden” o modelan la dinámica del fenómeno en estudio y tienen en cuenta el tiempo como una variable que afecta el fenómeno; de esta forma ofrecen resultados más robustos y cercanos a la realidad.

Uno de los puntos más importantes del procedimiento es el aprendizaje de la red neuronal. En términos simples lo que se hace es “sintonizar” o ajustar sus pesos de tal forma que el comportamiento global de la red sea lo más adecuado posible a la realidad.

Para el proceso de aprendizaje de una red neuronal se han utilizado diferentes algoritmos, cada uno con una serie de variantes que se desempeñan mejor que otros dependiendo de las condiciones y escenarios de cada problema. Uno de los que han sido más utilizados hasta la actualidad es el algoritmo *Back-propagation* (Cevalloa, 2004).

Recientemente se ha realizado estudios que evidencia que la utilización de los algoritmos genéticos para ajustar los pesos de una red neuronal ofrece resultados más exactos a la realidad que con la utilización de algoritmos tradicionales como el *Back-propagation*. Debe tomarse en cuenta que a pesar de que han sido más exactos, su aplicación es mucho más lenta.

La finalidad de las diferentes técnicas de pronósticos es obtener proyecciones lo más cercanas a la realidad; sin embargo, cada técnica funciona bajo ciertas circunstancias dadas o supuestos. La finalidad de utilizar las redes neuronales en este trabajo como herramienta de pronóstico de casos nuevos de pacientes hospitalizados y estancia hospitalaria, para pacientes con cáncer atendidos en los hospitales de la CCSS, es que pueda aprender con el paso del tiempo de las variaciones del comportamiento de los datos y se ajuste automáticamente al nuevo comportamiento, tomando en cuenta además otras variables independientes que puedan afectar los resultados esperados. Los algoritmos genéticos se utilizan para optimizar ese proceso de aprendizaje de la red en combinación con el método de propagación hacia atrás (*Back-propagation*).

## **1.2 Justificación del problema de investigación**

La necesidad de contar con pronósticos de demanda en las organizaciones, se puede visualizar en muchos sentidos, desde aspectos generales como la creación de planes estratégicos o la planeación de inversiones, hasta aspectos más específicos como el programa de producción, plan de requerimientos de materiales o equipos, o incluso para tomar decisiones como la necesidad de laborar horas extras de los equipos de trabajo en un día específico.

En el campo de la salud, es de fundamental importancia disponer de proyecciones efectivas de la demanda de los servicios de salud, con el fin de realizar una adecuada

planificación, garantizando la disposición oportuna de dichos servicios y una utilización eficiente de los recursos.

Una de las patologías prioritarias, que corresponde a una de las principales causas de muerte en el país, es el cáncer; razón por la cual la proyección de la demanda de los servicios utilizados por estos pacientes es clave para garantizar la atención oportuna de los mismos, buscando una mejor calidad de vida y sobrevida de estas personas.

Existen diversos tipos de pronósticos así como diferentes aplicaciones informáticas para la aplicación de los mismos. La mayoría de los modelos de pronósticos actuales utilizan supuestos o parámetros relacionados con la tendencia, estacionalidad y ciclicidad de los datos. Estos modelos han sido muy efectivos cuando se trabaja con datos que presentan un comportamiento uniforme a lo largo del tiempo, pero, en los casos en los que el comportamiento de los datos varía o que parecen no tener un comportamiento definido, lo ideal sería disponer de un modelo de pronósticos que “aprenda” de las variaciones y se ajuste automáticamente.

Las redes neuronales ofrecen la posibilidad de disponer de un modelo de pronósticos dinámico y no estático, gracias al aprendizaje que la red va teniendo a través del entrenamiento y aplicación, lo cual estaría potenciado con la utilización de los algoritmos genéticos en este proceso de entrenamiento de la red.

Para analizar los datos y hacer proyecciones de los pacientes nuevos hospitalizados con diagnóstico de cáncer en la CCSS, se requiere hacer uso de herramientas innovadoras como las redes neuronales y los algoritmos genéticos, con el fin de buscar resultados más cercanos a la realidad que los alcanzados con modelos tradicionales, de modo que permitan una mejor planificación de los servicios de salud.

Desde una perspectiva teórica esta investigación tiene una alta relevancia, ya que, se pretende combinar los pronósticos con dos técnicas muy novedosas en la investigación

administrativa, como lo son las redes neuronales y los algoritmos genéticos, esta relación podría derivar en soluciones óptimas para la planificación de la demanda de los servicios de salud de la CCSS y además podría abrir paso a la utilización de esta combinación de técnicas para afrontar otros problemas de la administración pública de la salud, mejorando los resultados obtenidos hasta el momento con las metodologías o los análisis utilizando técnicas tradicionales de optimización.

### **1.3 Formulación del problema de investigación**

¿Cuál es el aporte de las redes neuronales artificiales, optimizadas con algoritmos genéticos, en la consecución de proyecciones sobre los pacientes hospitalizados con cáncer atendidos en los hospitales de la CCSS, con resultados más precisos que los obtenidos con otras técnicas tradicionales de proyección?

### **1.4 Delimitaciones del estudio**

Con este proyecto se pretende ampliar el conocimiento teórico en relación con la utilización de las redes neuronales y los algoritmos genéticos en el desarrollo de modelos de proyección del número de pacientes con cáncer atendidos en hospitales de la CCSS y la estancia hospitalaria esperada para este tipo de paciente.

Dicho modelo debe contemplar variables como las características de los datos históricos, tendencia, estacionalidad, comportamientos cíclicos, características del entorno social y económico, entre otras variables, de manera que sea capaz de incluir tanto aspectos internos de la Institución y el tema de la salud, como externos.

Aunque estas técnicas se han desarrollado desde hace algunas décadas, su utilización en la optimización de procesos administrativos y organizacionales se ha iniciado en años más recientes, lo cual hace que este estudio se enfoque inicialmente en una investigación

exploratoria para determinar si es posible generar mejores pronósticos con redes neuronales que con los modelos estadísticos tradicionales.

Para realizar lo anterior se tomó una misma serie de datos y se le aplicaron las técnicas de pronósticos tradicionales y el modelo de redes neuronales desarrollado y se compararon los resultados con base en el error de pronóstico. La serie de datos analizada contempló datos disponibles en la CCSS sobre los casos nuevos de cáncer atendidos en hospitalización en los últimos 14 años.

Adicionalmente, se desarrolló un modelo para estimación de la estancia esperada de un paciente con cáncer hospitalizado en la CCSS, de manera que se pudiera realizar una mejor planificación de los servicios de internamiento de los hospitales y clínicas, de acuerdo con la situación de cada persona con esta enfermedad que llegue a ser atendida.

## **1.5 Objetivos**

### **1.5.1 Objetivo general**

Ampliar el conocimiento teórico en relación con la utilización de las redes neuronales y los algoritmos genéticos como modelos de proyección y optimización, con el fin de evaluar su aplicación práctica en la estimación del número de pacientes con cáncer hospitalizados en la CCSS y en la estimación de la estancia hospitalaria esperada para este tipo de pacientes.

### **1.5.2 Objetivos específicos**

1. Señalar los elementos teóricos y metodológicos requeridos para la construcción de un modelo de redes neuronales optimizadas con algoritmos genéticos.
2. Describir el comportamiento del número de pacientes con cáncer hospitalizados en la CCSS y de la estancia de este tipo de pacientes, con el fin de identificar las variables que deben considerarse en la elaboración de un modelo de redes neuronales para proyecciones sobre estos.

3. Diseñar un modelo para proyectar el número pacientes nuevos con cáncer hospitalizados en la CCSS, con el fin de evaluar la aplicación práctica de las redes neuronales artificiales como modelos de proyección.
4. Diseñar un modelo para estimar los días de estancia hospitalaria de los pacientes con cáncer atendidos en la CCSS, con el fin de evaluar la aplicación práctica de las redes neuronales artificiales como modelos de proyección.
5. Aplicar un procedimiento para la optimización de los modelos de redes neuronales diseñados, mediante algoritmos genéticos, con el fin de evaluar la posibilidad de disminuir el error de la red y mejorar las proyecciones.
6. Evaluar de forma comparativa el modelo de redes neuronales en relación con los modelos tradicionales de proyección de datos, con el fin de determinar cual método genera proyecciones con errores relativos más pequeños.
7. Elaborar una aplicación automatizada del modelo desarrollado que permita realizar de manera sencilla las proyecciones sobre los pacientes con cáncer, así como entrenar, optimizar y evaluar la efectividad del modelo periódicamente.
8. Identificar aportes teóricos y metodológicos en el diseño de modelos de redes neuronales optimizadas con algoritmos genéticos, con el fin de facilitar el desarrollo de futuras aplicaciones utilizando estas técnicas.

## **1.6 Hipótesis**

“Los modelos predictivos con redes neuronales artificiales optimizadas con algoritmos genéticos producen estimaciones con errores relativos más pequeños que los generados por modelos tradicionales”.

Para comprobar la hipótesis no se desarrolló una prueba estadística específica, sino que es una hipótesis intuitiva de trabajo, basada en la comparación del error porcentual medio absoluto (MAPE) utilizado en la evaluación de pronósticos.

## CAPÍTULO 2. MARCO REFERENCIAL

En el presente capítulo se desarrolla el marco referencial de la investigación. El mismo se divide en tres apartados: marco contextual, en el cual se mencionan las investigaciones nacionales e internacionales que se han desarrollado sobre el tema objeto de investigación, marco conceptual en el que se explican los principales conceptos relacionados con las temáticas de esta investigación y el marco teórico en el cual se mencionan los autores que han escrito sobre la temática y se define la perspectiva que tomará el grupo para el abordaje de este trabajo de investigación. Los dos primeros apartados se subdividen en los tres temas relacionados con esta tesis: pronósticos, redes neuronales y algoritmos genéticos.

### 2.1 Marco contextual

Dentro de la investigación realizada se incluyeron tres grandes temas relacionados, como lo son los Pronósticos de Variables, las Redes Neuronales y los Algoritmos Genéticos. Sin embargo, estos temas por sí solos han tenido un desarrollo particular, en donde han sido utilizados en diferentes investigaciones que son importantes de reseñar.

#### 2.1.1 Pronósticos

*“La fascinación del hombre ante el futuro se remonta a la era prehistórica. La gente siempre ha querido conocer el futuro por razones psicológicas (para combatir su miedo e inquietud ante lo desconocido)... Hay una base científica para pronosticar, pero también hay serias limitaciones en nuestra capacidad para predecir hechos y situaciones futuras. Es importante saber exactamente qué es lo que se puede y lo que no puede predecirse y las ventajas de los pronósticos así como sus limitaciones.” (Makridakis, 1992. Pág. 49).*

Con esta frase queda de manifiesto la necesidad del ser humano de anticipar las situaciones futuras que se le puedan presentar y con base en ello tomar medidas que le

ayuden a mejorar el aprovechamiento de las oportunidades o disminuir los perjuicios de las amenazas que se le presenten.

Desde la antigua Grecia, algunas instituciones se han dedicado a predecir el futuro como negocio, un ejemplo de ello fue el Templo de Delfos, quienes dirigieron el primer servicio institucionalizado de pronósticos (Makridakis, 1992). Con el paso del tiempo, la evolución del comercio, la creación de las empresas mercantiles y el desarrollo de todo tipo de organizaciones, hace que sea más necesario predecir las condiciones de diferentes variables, para lograr un adecuado desempeño.

Por su parte, el hecho de que el mundo en que se desenvuelven las organizaciones siempre ha sido cambiante y complejo, explica la necesidad que tienen de hacer pronósticos y previsiones futuras de diferentes variables. Adicionalmente, en los temas relacionados con la salud y la incidencia de ciertas enfermedades, la complejidad de una previsión aumenta, debido al efecto de diversas variables, tales como las condiciones propias de la población, la herencia genética, las condiciones sociales y económicas, que influyen en el estrés poblacional, los esfuerzos preventivos, entre otras.

Para mejorar los resultados de los pronósticos, se han creado una serie de técnicas, muchas de las cuales se empezaron a desarrollar y utilizar en el siglo XIX (Hanke, 1996), sin embargo, con la llegada de las computadoras se dio un impulso a la creación de técnicas más complejas, las cuales le permitieron a los administradores utilizar grandes volúmenes de datos para obtener mejores resultados y sin tener que dedicar notables esfuerzos en tiempo y recursos.

### **2.1.2 Redes Neuronales Artificiales**

Lo siguiente es una recapitulación del desarrollo teórico de las redes neuronales realizada por Hilera y Martínez (2000) es su obra *Redes Neuronales Artificiales: Fundamentos, modelos y aplicaciones*.

Los primeros teóricos en concebir los fundamentos de las redes neuronales fueron MacCulloch, un neurofisiólogo, y Pitts, un matemático, quienes en 1943 lanzaron una teoría acerca de la forma de trabajar de las neuronas. Ellos modelaron una red neuronal simple mediante circuitos eléctricos.

En la década de los cincuenta Marvin Minsky comenzó a construir la primera neurocomputadora, basada en los modelos de redes neuronales que imitaban al cerebro.

En 1959 Bernard Widrow y Marcial Hoff, de Stanford, desarrollaron el modelo *Adaline* (*Adaptive Linear Elements*). Esta fue la primera red neuronal aplicada a un problema real (filtros adaptativos para eliminar ecos en las líneas telefónicas).

En 1962 Frank Rosenblatt comenzó el desarrollo del *Perceptron*, el cual constituyó la red neuronal más antigua y la cual se utiliza hoy en día en el reconocimiento de patrones.

En 1967 Stephen Grossberg, de la Universidad de Boston, realizó la red *Avalancha*, que consistía en elementos discretos con actividad que varía con el tiempo, que satisface ecuaciones diferenciales continuas, para resolver actividades tales como reconocimiento continuo del habla y aprendizaje del conocimiento de los brazos de un robot.

En 1969 Marvin Minsky y Seymour Papert, del Instituto Tecnológico de Massachusetts (MIT) publicaron el libro *Perceptrons*, el cual además de contener el análisis matemático detallado del perceptron, consideraba que la extensión a perceptrones multinivel era completamente estéril. A partir de esta publicación surgieron numerosas críticas que ocasionaron que se detuviera la investigación en el tema y se le diera énfasis a otros temas relacionados, como la inteligencia artificial.

En 1977 James Anderson desarrolló un modelo lineal llamado *Asociador Lineal*; este modelo se basaba en el principio de que las conexiones entre neuronas son reforzadas cada vez que están activadas. Posteriormente Anderson desarrolló una potente extensión del asociador llamado *State in a Box*. Teuvo Kohonen, un ingeniero electrónico de la

universidad de Helsinki, desarrolló en Europa un modelo similar al de Anderson pero independiente.

En 1980 Kunihiro Fukushima definió el *Necognitron*, un modelo de red neuronal para el reconocimiento de patrones visuales.

En 1982 John Hopfield presentó su trabajo sobre redes neuronales ante la Academia Nacional de las Ciencias. En el trabajo describió a la red desarrollada a la cual le dio su mismo nombre, *Redes de Hopfield*, en las que utilizó funciones de energía para entender las redes dinámicas.

A la recapitulación anterior, Calvo (2004) agrega que Cohen y Grossberg desarrollaron en 1983 el principio de la memoria direccional, y en 1986 Rumulhart, Hunton y Williams redescubrieron el algoritmo de *Back-propagation* (desarrollado en 1974 por Paul Werbor) para el aprendizaje de redes neuronales.

### **2.1.3 Algoritmos genéticos (AG)**

Los AG fueron desarrollados por John Holland, sus colegas y sus estudiantes en la Universidad de Michigan durante la década de los setenta. Según Ramírez (1999), sus investigaciones tenían como objetivos: abstraer y explicar de manera rigurosa el proceso de adaptación de los sistemas naturales y diseñar sistemas artificiales de software que pudieran retener los mecanismos de los sistemas naturales.

Estos procedimientos se basaron en el reconocimiento de los procesos de adaptación a los que se han sometido los seres vivos a través de millones de años, lo cual hizo que pudieran sobrevivir en un mundo cambiante. Por tanto, de la misma manera se podría tener un conjunto de potenciales soluciones para un problema, las cuales se irían evaluando y seleccionando las mejores hasta que se obtuviera una solución óptima o al menos satisfactoria.

De acuerdo con Ramírez (1999), la principal característica que se ha investigado en los AG es la robustez, o sea, el balance entre la eficiencia y eficacia que generan las soluciones, dentro de los muchos y diferentes ambientes en que se utilicen. Esto proporciona ventajas en cuanto a la disminución de costos de rediseño y el mejoramiento de los resultados obtenidos.

Por su parte, Zúñiga (2004) señala que a diferencia de otros métodos de búsqueda y optimización, los AG tienen cuatro elementos que los caracterizan:

- a- Se aplica sobre muchas soluciones al mismo tiempo
- b- Evalúa directamente las funciones de costos, no depende de derivadas, funciones convexas u otros.
- c- Usa reglas probabilísticas para realizar los cambios en la población
- d- La codificación se basa en los parámetros de la solución y no en la solución misma.

#### **2.1.4 Investigaciones conjuntas**

En los últimos años, con el desarrollo más avanzado de las técnicas de los AG y las RNA, a nivel mundial se han realizado algunas investigaciones que relacionan estos dos temas, utilizándolos de manera conjunta para buscar obtener mejores resultados.

Orozco y Medina (2004) utilizaron la asociación de los AG y las RNA para crear un modelo de evolución del conocimiento en individuos en un sistema autoregulado, el cual permitía entender y manipular la evolución de las ideas y su transformación en conocimiento dentro de la mente humana, asimismo se pretendía investigar cómo hacer para que un colectivo de individuos pudiera transmitir sus conocimientos a otros; de tal manera que lo que se buscaba era cómo hacer que los organismos pudieran comenzar su vida con conocimientos preestablecidos que ya han sufrido un proceso de selección por sus antecesores y que por lo tanto fueron mejorados con el pasar del tiempo.

Otros investigadores han utilizado estas técnicas para crear robots que se adaptan al ambiente cambiante gracias a algoritmos evolutivos y controladores neuronales, tal es el caso de Nolfi (citado por Orozco y Medina [2004]), quien junto con su equipo desarrollaron Khepera, un robot con estas características, para el cual desarrolló un sistema de auto aprendizaje con redes neuronales artificiales, las cuales seleccionaban sus unidades de entrada mediante algoritmos genéticos.

Por otra parte, existen casos en nuestro país donde se han desarrollado aplicaciones que combinan los pronósticos con los algoritmos genéticos, tal es el caso de Calvo (2004) y de Pacheco, Sancho y Zoch (2006), quienes utilizaron los algoritmos genéticos para generar pronósticos de variables económicas y ventas, respectivamente; sin embargo, no se han presentado investigaciones completas a nivel nacional que relacionen las tres variables y donde se haya generado un modelo para pronosticar algún fenómeno o evento usando redes neuronales optimizadas con algoritmos genéticos, tampoco se han encontrado investigaciones para determinar casos nuevos de cáncer o de alguna otra enfermedad, utilizando AG, RNA o ambas combinadas, solamente han existido estudios de proyección de casos de cáncer, tanto para tipos específicos como en general, pero, estos fueron basados en proyecciones actuariales, realizados especialmente en la misma CCSS; por lo que esta investigación sería de las primeras que se realicen en nuestro país, combinando estos temas y para un campo como la predicción de enfermedades.

## **2.2 Marco Conceptual**

### **2.2.1 Pronósticos**

#### **2.2.1.1 Definición**

El concepto de pronóstico ha sido relacionado desde su creación con la estimación de niveles de ventas o demanda de productos o servicios, por lo que su conceptualización ha sido más enfocada en estos temas, que en otro tipo de aplicaciones o campos. Para la

revista VIRTUALPRO (N°104, Septiembre 2010), los pronósticos de demanda son la estimación de la cantidad de un producto o servicio que será adquirido por los consumidores. De ahí la importancia que poseen en la planificación de las acciones de la empresa, con vistas en satisfacer esa demanda proyectada o en su defecto, para buscar alternativas en caso de que esta demanda proyectada este por debajo de la capacidad de producción de la empresa.

A pesar de que esta definición está enfocada en el tema de ventas; la misma se puede interpretar perfectamente en el caso de proyección de incidencia hospitalaria del cáncer, dado que en este caso sería la estimación del número de casos nuevos de cáncer atendidos en los hospitales de la CCSS, lo cual se convierte en un insumo para que la Institución defina las acciones para atender esa demanda de servicios, en comparación con la capacidad de atención con que cuente en cierto momento.

De acuerdo con Camacho (2001), los pronósticos son proyecciones de los niveles futuros de venta de un sector o compañía en particular. Dentro de una empresa se pueden realizar a más de un nivel de acuerdo con el interés existente, con lo que es posible pronosticar las ventas totales de la empresa, de un grupo o familia de productos o de un artículo en particular; nuevamente trasladando estos conceptos al campo de aplicación en salud sobre la incidencia hospitalaria del cáncer, también se pueden hacer niveles y estimar la incidencia de algún tipo de cáncer en específico, o para una región o centro de salud en particular.

Está segregación a casos específicos ayudaría de gran manera a lograr una mejor focalización de los esfuerzos, planificación operativa y coordinación entre diferentes instancias dentro de la Institución, lo cual permitiría controlar mejor la enfermedad y mejorar la condición de los pacientes y la población en general.

Para Angelini (2000), el uso de modelos de pronósticos se ha ido extendiendo con el tiempo, sin embargo, es necesario lograr un equilibrio entre los costos y la precisión, ya

que, para obtener mayor precisión generalmente se necesita un modelo más sofisticado, lo cual va a incidir en mayores costos para ponerlo en práctica y mantenerlo en operación. No obstante, este mayor costo se suele compensar con el mayor ahorro en costos de operación al contar con mejor planificación de sus actividades.

#### **2.2.1.2 Aplicación**

Para Hanke (1996) cualquier organización, indiferentemente del tamaño o de si es pública o privada, utiliza el pronóstico ya sea explícito o implícito, debido a que casi todas las organizaciones deben planear como enfrentar las condiciones futuras de las cuales no tiene conocimiento cierto. Además, esta necesidad cruza todas las áreas funcionales, ya que, se requiere hacer pronósticos en las áreas de finanzas, recursos humanos y desarrollo de servicios. Por ejemplo, en el sector público, como el caso de la CCSS, al momento de crear los presupuestos de trabajo, se debe hacer uso de pronósticos para determinar los costos en que se va a incurrir, para satisfacer las necesidades de servicios estimados que requerirá la población.

#### **2.2.1.3 Análisis de series de datos**

La determinación del pronóstico se basa, en la mayoría de los casos, en los datos históricos que se tienen sobre la variable que se quiera pronosticar. Esto hace que sea necesario realizar un análisis de estos datos históricos, a fin de encontrar características que puedan ayudar a mejorar la precisión de pronóstico, tanto en la escogencia del método como en las variables que se incluyen en este. De acuerdo con Hanke (1996) los componentes de las series de tiempo que se deben considerar al momento de utilizar algún modelo son:

- **Tendencia:** La tendencia es el componente de largo plazo que representa el crecimiento o disminución en los valores de los datos de la serie, considerando un período largo de tiempo.

- **Estacionalidad:** Es el patrón de cambio en los datos de la serie que se repite a sí mismo en los diferentes años.
- **Componente cíclico:** Es la variación en forma de onda alrededor de la tendencia, afectada por lo regular por las condiciones económicas que se presentan.
- **Componente aleatorio:** Es el componente que mide la variabilidad en las series de tiempo, después de haber quitado los otros componentes, normalmente ocasionada por factores imprevistos y no recurrentes.

Angelini (2000) incluye un patrón más al que denomina horizontalidad, la cual se presenta cuando no existe una tendencia determinada, con lo que los datos aunque aumenten y disminuyan, en el largo plazo se tienden a mantener en los mismos rangos; este es un efecto que posiblemente sea observable en servicios como los que presta la Caja, pues si bien es cierto existe un componente poblacional, también es cierto que la capacidad instalada es limitada y en consecuencia aunque existe variación en la demanda la oferta se mantiene estable.

#### **2.2.1.4 Tipos de pronósticos**

De acuerdo con el período de tiempo abarcado, Hanke define dos tipos de pronósticos, los de largo plazo y los de corto plazo. Los de largo plazo son necesarios para establecer el curso general de la organización, por lo que son de uso particular de la alta dirección; mientras que los de corto plazo se utilizan para diseñar estrategias inmediatas, utilizadas por los mandos medios y de primera línea para enfrentar las necesidades operacionales inmediatas.

Otra clasificación que se puede realizar es de acuerdo con los mecanismos o técnicas para generarlos, según Edelman (*s.f.*) se podrían clasificar en: Técnicas de Proyección, Causalidad y Juicio Experto. Los basados en proyección suponen que el estudio del comportamiento pasado de la demanda permite detectar patrones que de alguna manera

sugieran el comportamiento futuro, para ello requieren el registro de datos históricos de la demanda y su posterior análisis cuantitativo.

Estos modelos se basan en las series de tiempo, es decir en un conjunto de observaciones de una variable a lo largo del tiempo (Monks, 1994, mencionado por Camacho, 2001). Según Camacho, estos métodos gozan de alto grado de confianza en mercados estables, donde no se presentan cambios importantes en la situación de la demanda del producto, la tecnología, competencia y entorno económico; por lo que son mayormente utilizados en previsiones de corto plazo. Sin embargo, cuando no se cuenta con un entorno estable y se espera que se produzcan cambios significativos, la proyección con base en los históricos no producirá resultados totalmente satisfactorios, por lo que es recomendable combinarlos con otros métodos como los de causalidad o juicio experto.

Los de causalidad se basan en construir modelos de causa-efecto de la variable buscada, respecto a otras variables de comportamiento conocido o previsible. Una vez demostradas tales relaciones, las estimaciones sobre las variables independientes conducen a una estimación de su efecto en la variable desconocida. El uso de datos históricos consistentes es esencial para comprobar la validez de los modelos establecidos. Finalmente, Edelman (*s.f.*) expresa que el juicio experto utiliza opiniones y apreciaciones subjetivas de expertos para componer una estimación de lo que puede suceder.

Dentro de los modelos de proyección de demanda se encuentran varios tipos, los cuales se pueden aplicar dependiendo de las características de la demanda del producto o servicio, algunos de estos modelos considerados tradicionales, por su habitual utilización, son: promedio simple, promedios móviles, suavización exponencial simple, suavización exponencial doble, Winters, Box-Jenkins.

**Promedio Simple (PS):** Para Angelini, este modelo utiliza datos históricos en donde los valores de todos los períodos ocurridos son ponderados con el mismo peso  $1/n$ ; donde  $n$  es el número de períodos.

$$PS = \sum \frac{\text{Valor de todos los períodos}}{\text{Número de períodos}}$$

**Promedio Ponderado (PP):** El mismo Angelini define que este modelo es semejante al promedio simple con la diferencia que se le asigna un peso diferente a cada uno de los períodos en la serie de tiempo.

$$PP = \sum \frac{\text{Valor de cada período} * \text{Peso de ponderación}}{\text{Número de períodos}}$$

Se debe tomar en cuenta que cada peso ponderados debe tener un valor entre 0 y 1 y que su sumatoria total no debe pasar de 1.

**Promedio Móvil Simple (PMS):** Este tipo de pronóstico, de acuerdo con Angelini, combina los datos de varios de los períodos más recientes, tomando ese promedio como el estimado para el próximo período. El número de períodos seleccionados debe permanecer constante de ese momento en adelante para ese mismo pronóstico. Una vez definido el número de períodos se ponderan los valores todos por igual. Este tipo de promedio se mueve, es decir una vez que se conoce un nuevo valor, se toma en cuenta esta y se descarta la del período más antiguo que se estaba tomando en cuenta.

$$PMS = \sum \frac{\text{Valor de los "n" últimos períodos}}{\text{Número de períodos en el promedio móvil}}$$

**Promedio Móvil Ponderado (PMP):** Según Angelini, este modelo incluye una ponderación para la demanda de los períodos pasados, diferente para cada uno de ellos. Utiliza el mismo procedimiento del promedio móvil simple, solamente que cada valor que se toma en cuenta se le debe multiplicar por el valor de ponderación.

$$PMP = \sum \frac{\text{Valor de los "n" últimos períodos} * \text{Peso de ponderación}}{\text{Número de períodos en el promedio móvil}}$$

Al igual que en el promedio ponderado, la suma de los factores no debe exceder de 1.

**Suavización exponencial simple (SES):** Para el Instituto Tecnológico y de Estudios Superiores de Monterrey (2006), estos modelos pronostican otorgando una ponderación a los datos dependiendo del peso que tengan dentro del cálculo del pronóstico. Esta ponderación se lleva a cabo a través de otorgarle un valor a la constante de suavización,  $\alpha$ , que puede ser mayor que cero pero menor que uno.

La suavización exponencial utiliza un promedio móvil ponderado de los datos históricos de la serie de tiempo como pronóstico. El modelo básico de suavización exponencial se presenta a continuación:

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$

Donde:

$F_{t+1}$  = Pronóstico simple de la serie de tiempo para el período de  $t + 1$ .

$Y_t$  = Valor real del período anterior al que se está pronosticando.

$F_t$  = Valor del pronóstico del período anterior al que se está pronosticando.

$\alpha$  = Constante de suavización ( $0 \leq \alpha \leq 1$ ).

Dentro de este modelo de pronóstico el valor de la constante de suavización posee una importancia alta, ya que, es a través de esta que se va ajustando el valor del pronóstico para irse acercando cada vez más a la realidad, de manera que se disminuya el error de la estimación con respecto a la realidad, que al final es lo que se busca al utilizar diferentes modelos.

**Suavización exponencial doble (SED):** Según Angelini, estos modelos están orientados a la eliminación de la distorsión en series de demanda estable. El modelo suaviza el pronóstico de la suavización simple y el pronóstico anterior de la suavización doble.

$$P_{t+1} = \alpha F_{t+1} + (1 - \alpha)P_t$$

Donde:

$P_{t+1}$  = Pronóstico doble de la serie de tiempo para el período de  $t + 1$ .

$F_{t+1}$  = Pronóstico simple de la serie de tiempo para el período de  $t + 1$ .

$P_t$  = Pronóstico doble del período anterior al que se está pronosticando.

$\alpha$  = Constante de suavización ( $0 \leq \alpha \leq 1$ ).

El poder de la suavización doble radica en que amortigua el “ruido”, al suavizar doblemente las demandas anteriores. Esta potencia hace al modelo conservador respecto al cambio, porque reacciona en forma lenta a los cambios en los patrones de demanda.

**Suavización exponencial Estacional: Modelo de Winters.** De acuerdo con el mismo Angelini, este modelo se utiliza para demandas que poseen tendencia y estacionalidades, es decir que se acomodan a condiciones artificialmente creadas por el hombre.

El modelo básico es:

$$D = (a + b_t)I_t + E_t$$

Donde:

D = Valor pronosticado

a = Nivel

b = Tendencia

$I_t$  = Índice estacional apropiado para el período t

$E_t$  = Variables aleatorias independientes con media  $\mu$  y varianza  $\sigma^2$ .

El uso repetido de este modelo necesita estar actualizando las variables a, b, e  $I_t$ , para lo cual Winters en 1960 sugirió el uso de las siguientes ecuaciones:

$$\hat{a}_t = \alpha \left( \frac{D_t}{I_{t-p}} \right) + (1 - \alpha)(\hat{a}_{t-1} + b_{t-1})$$

$$b_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1 - \beta)(b_{t-1})$$

$$I_t = \gamma \left( \frac{D_t}{\hat{a}_t} \right) + (1 - \gamma)(\hat{I}_{t-p})$$

Donde:

$\alpha$ ,  $\beta$ ,  $\gamma$ , son las tres constantes de suavización, las cuales tienen un valor entre 0 y 1.

Además el término  $(D_t / I_{t-p})$ , refleja un estimado de la estacionalidad deducida del valor real en el período t, en el cual  $I_{t-p}$  fue el estimado de índice estacional para los más recientes períodos equivalentes en el ciclo estacional. Los restantes términos de las ecuaciones son equivalentes a los no estacionales.

Por su parte el término  $(D_t / \hat{a}_t)$ , provee un estimado del factor estacional basado en las últimas observaciones obtenidas, por lo que se refleja una combinación del estimado histórico con el estimado basado en los últimos datos.

**Suavización exponencial ajustada a la tendencia: Método de Holt.** Según Hanke (1996) este modelo atenúa en forma directa la tendencia y la pendiente empleando diferentes constantes de suavización para cada una de ellas. Esta técnica proporciona mayor flexibilidad al seleccionar las proporciones a las que se rastrearán la tendencia y la pendiente. Las ecuaciones utilizadas son:

$$A_t = \alpha Y_t + (1 - \alpha)(A_{t-1} + T_{t-1})$$

$$T_t = \beta(A_t - A_{t-1}) + (1 - \beta)T_{t-1}$$

$$\hat{Y}_{t+p} = A_t + pT_t$$

Donde:

$A_t$  = Nuevo valor suavizado

$\alpha$  = constante de suavización de los datos ( $0 \leq \alpha \leq 1$ )

$Y_t$  = nueva observación o valor real de la serie en el período t

$\beta$  = constante de suavización de la estimación de la tendencia ( $0 \leq \beta \leq 1$ )

$T_t$  = Estimación de la tendencia

$p$  = períodos a pronosticar en el futuro

$\hat{Y}_{t+p}$  = pronóstico de p períodos en el futuro.

**Método de Box-Jenkins:** Para Hanke, este método de pronóstico es diferente a la mayoría, ya que, no asume ningún patrón particular en los datos históricos, al contrario utiliza un enfoque iterativo de identificación de un modelo útil a partir de los modelos generales.

La metodología consiste en tres etapas básicas:

*Etapa 1. Identificación del modelo:* Primeramente se determina si la serie es estacionaria; o sea, si el valor de la media varía a través del tiempo. En caso de que no lo sea, se puede convertir mediante el método de diferenciación. Posterior a la obtención de la serie estacionaria, el analista debe identificar la forma del modelo a utilizar.

*Etapa 2. Estimación del modelo y prueba de su adecuación:* Una vez estimado el modelo tentativo, se deben estimar los parámetros para ese modelo. Ya con los parámetros se debe verificar si este es adecuado, para lo cual se deben revisar los términos de error y asegurarse de que son aleatorios. Esta verificación se puede hacer revisando que las autocorrelaciones de los errores no sean diferentes de cero en forma significativa, en caso de que se presenten autocorrelaciones muy diferentes de cero, el modelo no es el adecuado y se debe buscar otro.

*Etapa 3. Pronóstico con el modelo:* Una vez que se encontró el modelo adecuado, se procede con la realización del pronóstico para los períodos que se requiera. En esta etapa, si se encuentran pequeñas diferencias en los errores de pronóstico, podría necesitarse de una modificación en los parámetros empleados en el modelo; sin embargo, si estas diferencias son muy grandes, es posible que se necesite un modelo completamente nuevo.

#### **2.2.1.5 Error de pronóstico**

De acuerdo con Ramírez (2004), el error de pronóstico se refiere a la diferencia entre el valor pronosticado y el valor real de una variable dada, para un período específico. El error de pronóstico sirve para determinar el mejor método de pronóstico para una serie de tiempo, analizando las diferencias presentadas en varios períodos. Existen 3 principales métodos de medición del error de pronóstico: el MAD, MSE y el MAPE.

El MAD (Mean Absolute Deviation – Desviación Media Absoluta): Esta consiste en la sumatoria de los errores absolutos (diferencia entre el valor real y el pronóstico) dividido entre el número de períodos incluido en la sumatoria.

$$MAD = \frac{\sum |e|}{n}$$

El MSE (Square Error – Error Medio Cuadrado) consiste en la sumatoria de los errores de pronóstico al cuadrado, dividido por el número de períodos incluidos.

$$MSE = \frac{\sum e^2}{n}$$

Por su parte el MAPE (Mean Absolute Porcentual Error – Error Porcentual Medio Absoluto) es la sumatoria de las diferencias porcentuales entre los valores reales y el pronóstico (PE), medidas en función del valor real, dividido por el número de períodos utilizados en la suma.

$$MAPE = \frac{\sum PE}{n}$$

## 2.2.2 Redes Neuronales Artificiales

### 2.2.2.1 Definición

Las redes neuronales artificiales son modelos matemáticos que tratan de hacer una pequeña emulación del funcionamiento del cerebro humano. Estas redes se aprenden la dinámica del fenómeno en estudio y tienen en cuenta el tiempo como una variable que afecta el fenómeno. Las siguientes son definiciones del concepto de redes neuronales artificiales (RNA), según diferentes autores.

García, citado por Gallardo (2009), indica que las redes neuronales artificiales (RNA) *“son sistemas adaptivos que aprenden a partir de su propia experiencia”* (p. 12).

Para Calvo (2004) *“las redes neuronales son modelos matemáticos multivariantes que utilizan procedimientos iterativos, en general minimizan funciones de error”* (p. 6).

Kohonen, citado por Hilera y Martínez (2000), describe las redes neuronales como redes interconectadas masivamente en paralelo de elementos simples (usualmente adaptativos) y con organización jerárquica, las cuales intentan interactuar con los objetos del mundo real del mismo modo que lo hace el sistema nervioso biológico.

De acuerdo con Solano (1997), la RNA es una red de procesadores o unidades muy simples (cada una con posibilidad de tener una pequeña cantidad de memoria local), y masivamente conectadas entre sí por enlaces de comunicación unidireccional (denominado aquí como conexiones) que llevan un valor numérico (en lugar de datos simbólicos), correspondiente al peso de la conexión entre unidades.

*“Una red neuronal es un sistema de procesadores paralelos conectados entre sí en forma de grafo dirigido. Esquemáticamente cada elemento de procesamiento (neuronas) de la red se representa como un nodo. Estas*

*conexiones establecen una estructura jerárquica que tratando de emular la fisiología del cerebro busca nuevos modelos de procesamiento para solucionar problemas concretos del mundo real” (Freman y Skapura, citados por Sotolongo y Guzmán, 2001, p. 3).*

*“Las RNA son modelos matemáticos que tratan de hacer una pequeña emulación del funcionamiento del cerebro humano” (Toro, Mejía y Salazar, 2004, p.26).*

Finalmente, de acuerdo con Nojek, Britos, Rossi y García (2003) las redes neuronales son sistemas dinámicos autoadaptativos. Son adaptables debido a la capacidad de autoajustarse de los elementos procesales (neuronal) que componen el sistema. Son dinámicos, pues son capaces de estar constantemente cambiando para adaptarse a las nuevas condiciones.

#### **2.2.2.2 Ventajas**

Debido a su constitución y a sus fundamentos, las redes neuronales artificiales presentan un gran número de características semejantes a las del cerebro. Esto hace que ofrezcan numerosas ventajas y que este tipo de tecnología se esté aplicando en múltiples áreas. Hilera y Martínez (2000) identifican las siguientes ventajas:

- Aprendizaje adaptativo: capacidad de aprender y realizar tareas basadas en un entrenamiento o una experiencia inicial.
- Autoorganización: una red neuronal puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje.
- Tolerancia a fallos: la destrucción parcial de una red conduce a una degradación de su estructura; sin embargo, algunas capacidades de las redes se pueden retener, incluso sufriendo un gran daño.
- Operación en tiempo real: las redes neuronales permiten realizar grandes procesos con datos de forma muy rápida, debido a su implementación paralela. Esta

propiedad hace que las redes neuronales sean la mejor alternativa para reconocimiento y clasificación de patrones en tiempo real.

- Fácil inserción dentro de la tecnología existente: las redes neuronales pueden ser entrenadas rápidamente, comprobadas, verificadas y trasladadas a una implementación de bajo costo.

### 2.2.2.3 Aplicaciones

Las redes neuronales son una tecnología emergente que puede utilizarse en un gran número y variedad de aplicaciones.

De acuerdo con Gallardo (2009) las RNA han resultado ser muy eficientes en la resolución de problemas de reconocimiento de patrones, tales como el reconocimiento de la voz (algoritmo *NETTALK*) y el de caracteres postales (como lo es el algoritmo de *Le Cun*).

Por su parte, Calvo (2004) indica que se pueden encontrar modelos de redes neuronales en programas de reconocimiento de voz, en juegos de ordenador, programas de contabilidad, tutores y muchos otros.

Algunas aplicaciones de las redes neuronales se relacionan con: procesamientos de imágenes y voz, reconocimiento de patrones, planeamiento, predicción, control y optimización (Toro, Mejía y Salazar, 2004).

En el siguiente cuadro se presentan las principales aplicaciones de las RNA en diferentes campos.

**Cuadro 1: Algunas aplicaciones de las RNA**

Área	Aplicación
Biología	a) Aprender más acerca del cerebro y otros sistemas, b) Obtención de modelos de la retina.
Empresa	a) Evaluación de probabilidad de formaciones geológicas y petrolíferas, b)

Área	Aplicación
	identificación de candidatos para posiciones específicas, c) explotación de bases de datos, d) optimización de plazas y horarios en líneas de vuelo, e) reconocimiento de caracteres escritos.
Medio ambiente	a) Analizar tendencias y patrones, b) previsión del tiempo.
Finanzas	a) Previsión de la evolución de los precios, b) valoración del riesgo de los créditos, c) identificación de falsificaciones, d) interpretación de firmas.
Manufactura	a) Robots automatizados y sistemas de control (visión artificial y sensores de presión, temperatura, gas, etc.), b) control de producción en líneas de proceso, c) inspección de la calidad.
Medicina	a) Analizadores del habla para la ayuda de audición de sordos profundos, b) diagnóstico y tratamiento a partir de síntomas y/o datos analíticos (electrocardiograma, encefalograma, análisis sanguíneo, etc.), c) monitorización en cirugía, d) predicción de reacciones adversas a los medicamentos, e) lectores de rayos X, f) entendimiento de la causa de los ataques epilépticos.
Militares	a) Clasificación de las señales de radar, b) creación de armas inteligentes, c) optimización del uso de recursos escasos, e) reconocimiento y seguimiento en el tiro al blanco, f) interfaces adaptativas para sistemas hombre/máquinas.

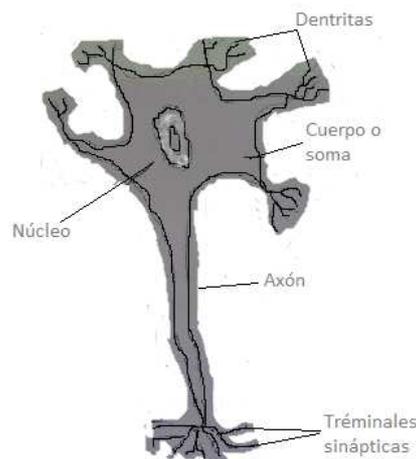
Fuente: Hilera y Martínez (2000, p. 21-22).

#### 2.2.2.4 Estructura y funcionamiento de las neuronas biológicas

El sistema nervioso humano está compuesto por el sistema nervioso central y el sistema nervioso periférico. El primero está conformado por la médula espinal y el encéfalo, en el cual se encuentra el cerebro. Dentro del cerebro se encuentran los nervios, los cuales nacen a partir del sistema nervioso central y recorren todo el cuerpo llevando las señales emitidas por éste. La unidad básica del sistema nervioso y del cerebro son las células llamadas neuronas. Una neurona está conformada básicamente por tres partes: el cuerpo de la célula o soma, las dendritas y el axón (Figura 1).

La mayoría de las neuronas reciben información de otras neuronas por medio de las dendritas y llevan información a otras por medio del axón, el cual conduce electricidad proveniente del mensaje enviado por el soma de la neurona. El proceso de transmisión de información de una neurona a otra se conoce con el nombre de sinapsis. Al ocurrir la sinapsis como consecuencia de la señal eléctrica proveniente del axón se liberan unas sustancias llamadas neurotransmisores, las cuales hacen posible el paso de la actividad eléctrica de una neurona a otra. Cuando las señales eléctricas sobrepasan un cierto umbral se dice que la neurona está en estado de excitación, una neurona transmite señales a las neuronas que se encuentran a su alrededor (Andrews, Carpenter, Rusell y Norving, citados por Gallardo, 2009).

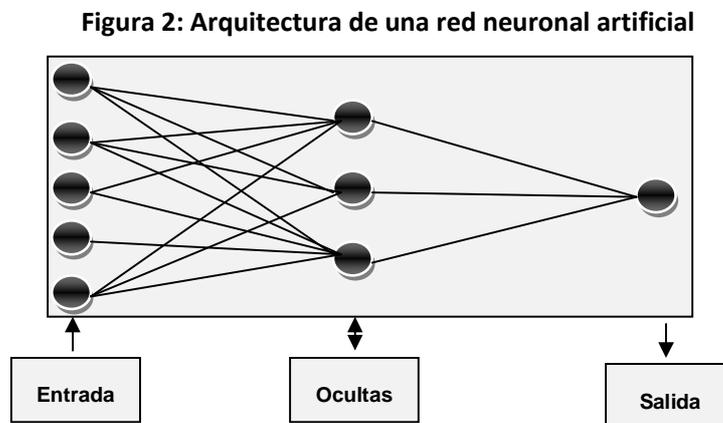
**Figura 1: Partes de una neurona biológica**



Fuente: Ugalde (1997, p.3).

En el caso de las redes neuronales artificiales, cada nodo representa una neurona, donde las dendritas son el medio por donde entra la información y el axón es la salida y además el responsable de excitar a otras neuronas, lo cual es reflejado en las RNA mediante funciones matemáticas, siendo la función de activación análoga a la acción de los neurotransmisores de las redes neuronales. En las RNA también se da un aprendizaje con

el fin de establecer el patrón entre las entradas y las salidas del modelo, este patrón llega a establecerse por medio de una regla de aprendizaje.



Fuente: Gallardo (2009, p.13)

#### 2.2.2.5 Estructura y funcionamiento de las RNA

La neurona artificial pretende imitar las características más importantes de las neuronas biológicas. Cada neurona  $i$ -ésima está caracterizada en cualquier instante por un valor numérico denominado valor o estado de activación  $a_i(t)$  asociado a cada unidad, existe una función de salida  $f_i$  que transforma el estado actual de activación en una señal de salida  $y_i$ . Dicha señal es enviada a través de los canales de comunicación unidireccionales a otras unidades de la red; en estos canales la señal se modifica de acuerdo con la sinapsis (el peso  $w_{ji}$ ) asociada a cada uno de ellos según una determinada regla. Las señales moduladas que han llegado a la unidad  $i$ -ésima se combinan entre ellas, generando así la entrada total,  $Net_i$ .

$$Net_j = \sum_i y_i w_{ji}$$

Una función de activación,  $f$ , determina el nuevo estado de activación  $a_i(t+1)$  de la neurona, teniendo en cuenta la entrada total calculada y el anterior estado de activación  $a_i(t)$ .

La dinámica que rige la actualización de los estados de las unidades (evolución de la red neural) puede ser de dos tipos: modo asincrónico y modo síncrono. En el primer caso, las neuronas evalúan su estado continuamente, según les va llegando información, y lo hacen de forma independiente. En el caso síncrono, la información también llega de forma continua, pero los cambios se realizan simultáneamente.

Solano (1997), Hilera y Martínez (2000) y Gallardo (2009) coinciden en que existen siete aspectos importantes sobre los que se basa la arquitectura de una RNA: un conjunto de unidades de procesamiento, un estado de activación, un patrón de conectividad, una función de activación, una función de salida, una regla de aprendizaje y una regla de propagación. Ugalde (1997) agrega un octavo componente: ambiente.

### **Unidades de procesamiento**

Se le llama arquitectura de una red neuronal a la estructura de interconexión del conjunto de nodos que componen la red, y que se organizan en capas.

Se conoce como capa o nivel a un conjunto de neuronas cuyas entradas provienen de la misma fuente (que puede ser otra capa de neuronas) y cuyas salidas se dirigen al mismo destino (que puede ser otra capa de neuronas).

Para el caso de los modelos más simples de RNA se puede encontrar una capa de entrada (*input*), y una capa de salida (*output*). Conforme fue avanzando el estudio de las RNA se vio la necesidad de incorporar capas intermedias, también llamadas ocultas, las cuales procesan la información dentro del sistema, pero, no son visibles desde fuera del mismo.

El número de nodos en la capa de entrada depende de las entradas consideradas para establecer un patrón entre la representación de la entrada y la de la salida. El número de nodos en la capa intermedia varía estableciéndose en muchos casos mediante prueba y error para elegir el número más conveniente. El número de nodos en la capa de salida depende del número de categorías que tenga la misma. Dentro de cada capa se incluye

una función de activación o de transferencia, la cual debe satisfacer algunas especificaciones necesarias para resolver el problema (Gallardo, 2009).

### **Estados de activación**

De acuerdo con Hilera y Martínez (2000) todas las neuronas que componen la red se hallan en cierto estado. En una visión simplificada, se puede decir que hay dos posibles estados, reposo y excitado, a los que se denominan globalmente estados de activación, y a cada uno de los cuales se le asigna un valor. Los valores de activación pueden ser continuos o discretos. Además, pueden ser limitados o ilimitados. Si son discretos, suelen tomar un conjunto pequeño de valores o bien valores binarios. En notación binaria, un estado activo se indicaría por un 1, y se caracteriza por la emisión de un impulso por parte de la neurona (potencial de acción), mientras que un estado pasivo se indicaría por un 0, y significaría que la neurona está en reposo. En otros modelos se considera un conjunto continuo de estados de activación, en lugar de sólo dos estados, en cuyo caso se les asigna un valor entre  $[0,1]$  o en el intervalo  $[-1,1]$ .

### **Patrón de conectividad**

Una conexión es una línea única de comunicación que va de una unidad de envío a una unidad receptora. Hay dos tipos de conexiones: excitación, que tienden a causar el disparo (activación) de las unidades (representado por un valor positivo), e inhibidores, que tienden a prevenir el disparo (desactivación) de la unidad (representado por valores negativos).

El patrón de conectividad entre unidades indica como estas unidades están conectadas entre sí y el tipo de conexión utilizado. Las conexiones que unen a las neuronas que forman una red neuronal artificial tienen asociado un peso, que es el que hace que la red adquiera conocimiento. Este peso controla la fuerza de las señales de entrada a la unidad, y se denota por  $w_{ji}$ , donde  $i$  representa la unidad receptora y  $j$  la unidad emisora.

Suele utilizarse una matriz  $W$  con todos los pesos  $w_{ji}$  que reflejan la influencia que sobre la neurona  $j$  tiene la neurona  $i$ .  $W$  es un conjunto de elementos positivos, negativos o nulos. Si  $w_{ji}$  es positivo, indica que la interacción entre las neuronas  $i$  y  $j$  es excitadora; es decir, siempre que la neurona  $i$  esté activada, la neurona  $j$  recibirá una señal de  $i$  que tenderá a activarla. Si  $w_{ji}$  es negativo, la sinapsis será inhibitoria. En este caso, si  $i$  está activada, enviará una señal a  $j$  que tenderá a desactivar a ésta. Finalmente, si  $w_{ji} = 0$ , se supone que no hay conexión entre ambas.

### **Regla de propagación**

Considérese  $y_i$  como el valor de salida de una neurona  $i$  en un instante dado. Una neurona recibe un conjunto de señales que le dan información del estado de activación de todas las neuronas con las que se encuentra conectada. Cada conexión (sinapsis) entre la neurona  $i$  y la neurona  $j$  está ponderada por un peso  $w_{ji}$ . Normalmente, como simplificación, se considera que el efecto de cada señal es aditivo, de tal forma que la entrada neta que recibe una neurona (potencial postsináptico)  $Net_j$  es la suma del producto de cada señal individual por el valor de la sinapsis que conecta ambas neuronas:

$$Net_j = \sum_i w_{ji} y_i$$

Esta regla muestra el procedimiento a seguir para combinar los valores de entrada a una unidad con los pesos de las conexiones que llegan a esa unidad y es conocida como regla de propagación.

Esta regla establece la forma en la cual la información se extiende a la RNA. La regla toma el vector de salida  $y(t)$  y lo combina con las matrices de conectividad  $w(t)$  para producir una red para cada tipo de entrada a la RNA.

En un modelo simple, la regla de propagación puede ser considerada como el patrón de conectividad.

## Función de activación

Así como es necesaria una regla que combine las entradas a una neurona con los pesos de las conexiones, también se requiere una regla que combine las entradas con el estado actual de la neurona para producir un nuevo estado de activación. Esta función  $f$  produce un nuevo estado de activación en una neurona a partir del estado ( $a_i$ ) que existía y la combinación de las entradas con los pesos de las conexiones ( $Net_i$ ).

Dado el estado de activación  $a_i(t)$  de la unidad  $u_i$  y la entrada total que llega a ella  $Net_i$ , el estado de activación siguiente,  $a_i(t+1)$ , se obtiene aplicando una función  $f$ , llamada función de activación.

$$a_i(t+1) = f(a_i(t), Net_i)$$

En la mayoría de los casos,  $f$  es la función identidad, por lo que el estado de activación de una neurona en  $t+1$  coincidirá con el  $Net$  de la misma en  $t$ . En este caso, el parámetro que se le pasa a la función de salida,  $f$ , de la neurona será directamente el  $Net$ ; el estado de activación anterior no se tiene en cuenta. Según esto, la salida ( $y_i$ ) de una neurona  $i$  quedará según la expresión:

$$y_i(t+1) = f(Net_i) = f\left(\sum_{j=1}^n w_{ij}y_j(t)\right)$$

Por tanto, se puede considerar la función  $f$ , indistintamente como función de transferencia o de activación. Además, normalmente la función de activación no está centrada en el origen del eje que representa el valor de la entrada neta, sino que existe cierto desplazamiento debido a las características internas de la propia neurona, ya que, no es igual en todas ellas. Este valor se denota como  $\theta_i$  y representa el umbral de activación de la neurona  $i$ .

$$y_i(t+1) = f(Net_i - \theta_i) = f\left(\sum_{j=1}^n w_{ij}y_j(t) - \theta_i\right)$$

## Funciones de salida

Entre las unidades o neuronas que forman una red neuronal artificial existe un conjunto de conexiones que unen unas a otras. Cada unidad transmite señales a aquellas que están conectadas con su salida. Asociada con cada unidad hay una función de salida que transforma el estado actual de activación en una señal de salida.

De acuerdo con Gallardo (2009) existen cuatro funciones de transferencia típicas que determinan distintos tipos de neuronas: función escalón, función lineal y mixta, función sigmoideal y función gaussiana.

Función escalón (*hardlim*). Cuando se utiliza este tipo de función en los modelos de RNA se obtienen salidas con valores de 1 ó 0, lo cual va a depender de si el valor sináptico es positivo o negativo (Figura 3). Esta función se utiliza en modelos de redes neuronales simples como por ejemplo el perceptrón monocapa. Esta función es de la forma:

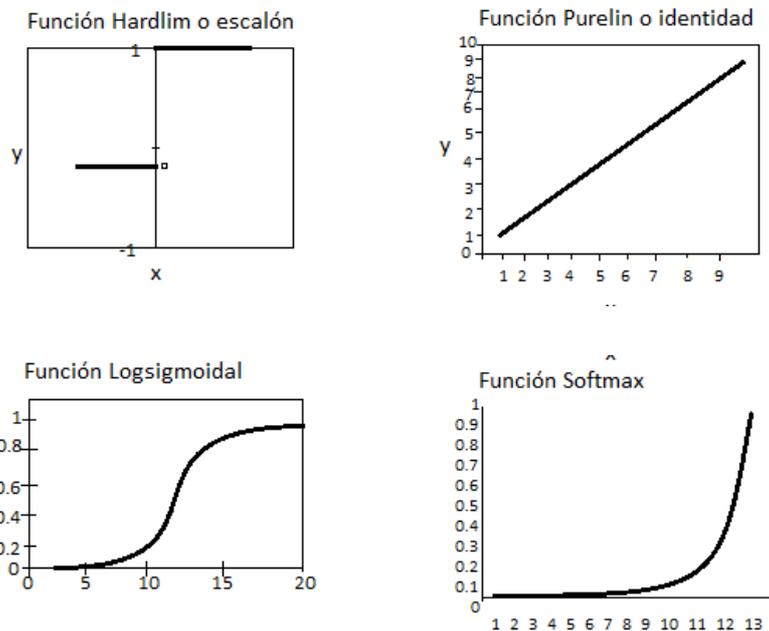
$$y = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$$

Función identidad (*purelin*). El nivel de activación es pasado directamente como salida. Se usan en diferentes tipos de modelos de RNA, incluidas las redes lineales y las salidas de las capas de la red función de base radial. Esta función es de la forma  $y = x$  (Ver figura 3).

Función sigmoideal. Se trata de una curva con forma de s (sigmoide), siendo el rango de salida un valor entre 0 y 1 (Figura 3). Su forma es la siguiente:  $y = \frac{1}{1 + e^{-x}}$

Función Softmax. Es una función exponencial, con resultados normalizados de modo que la suma de las activaciones a través de la capa es 1. Puede utilizarse en la capa de salida de perceptrón multicapa, para problemas de clasificación de más de dos clases, de modo que las salidas puedan ser interpretadas como probabilidades de membrecía. Su forma es la siguiente:  $y = \frac{e^{x_i}}{\sum_i e^{x_i}}$  (Ver Figura 3).

**Figura 3: Funciones de transferencia utilizadas en los modelos de Redes Neuronales Artificiales**



Fuente: Gallardo (2009, p.15).

### Regla de aprendizaje

El proceso de aprendizaje consiste en hallar los pesos que codifican los conocimientos. Una regla de aprendizaje hace variar el valor de los pesos de una red hasta que estos adoptan un valor constante ( $dw/dt=0$ ); cuando esto ocurre se dice que la red ya ha aprendido (Sotolongo y Guzmán, 2001).

Al igual que el funcionamiento de una red depende del número de neuronas de las que disponga y de cómo estén conectadas entre sí, cada modelo dispone de sus propias técnicas de aprendizaje. Virtualmente todas las reglas de aprendizaje para un modelo de RNA pueden ser consideradas como una variante de la regla de aprendizaje de Hebb. La versión más simple de la regla de Hebb es:

$$\Delta w_{ij} = \alpha a_i o_j$$

donde  $\alpha$  es la constante de proporcionalidad que representa la tasa de aprendizaje.

### Tipos de aprendizaje

Las RNA deben utilizar una regla de aprendizaje o algoritmo de entrenamiento, es decir, un procedimiento para entrenar a la red para que esta realice una determinada tarea. De acuerdo con Gallardo (2009) las reglas de aprendizaje se pueden clasificar de acuerdo a tres categorías:

1. Aprendizaje supervisado: para utilizar este tipo de aprendizaje es necesario tener un conjunto de ejemplos mediante los cuales la red aprende el patrón de los datos. Además es necesario tener la capa de entrada y la salida correcta; es decir, el patrón de los datos se establece a partir de los datos de entrada y la salida correcta de la red. Básicamente se trata del problema estadístico de discriminación.
2. Aprendizaje no supervisado: las entradas y salida son modificadas con base en la entrada, extrayendo características de los conglomerados que se conformen. Básicamente se trata del problema estadístico de generar conglomerados.
3. Aprendizaje reforzado o por reforzamiento: se encuentra entre el aprendizaje supervisado y el no supervisado con la novedad de que el aprendizaje se da externamente y son favorecidas aquellas acciones dirigidas al éxito y desfavorecidas aquellas dirigidas al fracaso.

### **Ambiente**

El entorno en el que el sistema va a funcionar es representada por un conjunto de patrones de entrada (vectores) que van a ser aprendidos por la RNA.

#### **2.2.2.6 Proceso de trabajo de las redes neuronales**

De acuerdo con Calvo (2004), el proceso de trabajo de las redes neuronales se detalla en los siguientes siete pasos.

1. El punto de partida lo constituye la incorporación de un conjunto de datos.
2. Después se desarrolla la arquitectura neuronal, es decir, se selecciona el modelo y tipo de red, el número de variables dependientes e independientes, el número de capas y el número de neuronas.
3. Se produce a la fase de aprendizaje o estimación del modelo. Una vez desarrollada la arquitectura, el conjunto de datos es aplicado al modelo para que la red aprenda a identificar las características propias del fenómeno en estudio.
4. Terminado el proceso anterior, con el modelo entrenado, se ejecuta una segunda prueba o fase de test para calificar de mejor forma la red y queda preparada para la siguiente fase.
5. Se observa la validez de los datos. Esto se logra al comparar los resultados de las fases de entrenamiento y test con los resultados reales que se deberían obtener. Esta comparación se hace necesaria ya que, es la manera de aceptar o rechazar el modelo vía cuantificación del error de predicción obtenido versus el límite permitido por los investigadores.
6. Si la cuantificación de error no es aceptable se realizan ajustes mediante nuevos patrones de test que calibran la red hasta que el error obtenido sea aceptable.
7. Una vez hecha la calibración y obtenido un modelo que refleje adecuadamente el fenómeno, la red queda preparada para entrar en operación.

#### **2.2.2.7 Tipos**

Según el tipo de arquitectura y funciones que se utilicen, existen diversos modelos de redes neuronales artificiales. Los modelos más conocidos se clasifican en supervisados y no supervisados. Dentro de los primeros se tienen los modelos más simples y las que más se utilizan: *Perceptrón*, *Back-propagation BP*, *Adalina*, *LVQ*, *GRNN*, *Perceptrón Multicapa*, *Madalina*. Dentro de los no supervisados los más comunes son el de Realimentación de Hopfield y el Unidireccional de Mapas de Kohonen. En general, el modelo más utilizado en

la actualidad es el de la Red *Back-propagation*, con diversos algoritmos de entrenamiento (Cevalloa, 2004).

La aplicación de redes neuronales que será objeto de la presente tesis es la correspondiente a predicción, ya que, la finalidad es desarrollar un modelo de pronósticos. En este sentido, el tipo de red neuronal que se ha utilizado con mayor éxito en la predicción es la red *Back-propagation* (Propagación hacia atrás), la cual se describirá a continuación.

#### **2.2.2.8 Red neuronal de propagación hacia atrás (*Back-propagation*)**

El algoritmo de propagación hacia atrás, o retropropagación fue desarrollado por Rumelhart y McClelland (Universidad de California en San Diego), en 1986. Este algoritmo se puede aplicar en modelos de redes con más de dos capas de células y permite la representación interna del conocimiento que es capaz de organizar en la capa intermedia de las células para conseguir cualquier correspondencia entre la entrada y la salida de la red.

En una red de propagación hacia atrás todas las unidades de proceso interfieren en una disminución del error (no solo las de salidas), por medio de la modificación de sus pesos; de ahí que el trabajo es dividido en muchas pequeñas tareas que cada neurona realiza.

De forma simplificada, el funcionamiento de esta red consiste en un aprendizaje de un conjunto predefinido de pares de entradas-salidas dados como ejemplo, empleando un ciclo propagación-adaptación de dos fases:

- Primero se aplica un patrón de entrada como estímulo para la primera capa de las neuronas de la red; este se va propagando a través de todas las capas superiores hasta generar una salida; luego se compara el resultado obtenido en las neuronas de salida con la salida que se desea obtener y se calcula un valor del error para cada neurona de salida.

- A continuación, estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida, recibiendo el porcentaje de error aproximado a la participación de la neurona intermedia en la salida original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total. Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se presente el mismo patrón, la salida esté más cercana a la deseada; es decir, el error disminuya.

De acuerdo con Hilera y Martínez (2000), la finalidad del algoritmo *Back-propagation* es determinar el ajuste necesario en los pesos de la red neuronal para disminuir el error, para lo cual se utilizan las siguientes fórmulas.

$$w_{ji}(t + 1) = w_{ji}(t) + [\Delta w_{ji}(t + 1)]$$

$$w_{ji}(t + 1) = w_{ji}(t) + [\alpha \delta_{pj} y_{pi} + \beta \Delta w_{ji}(t)]$$

Donde:

$$\delta_{pj} = (d_{pj} - y_{pi}) * f'(net_j) ; \text{ si } U_j \text{ es una neurona de salida.}$$

$$\delta_{pj} = (\sum_k \delta_{pk} w_{kj}) * f'(net_j) ; \text{ si } U_j \text{ no es una neurona de salida.}$$

De acuerdo con Freeman y Skapura (1993) e Hilera y Martínez (2000) los siguientes son los pasos y fórmulas a utilizar para aplicar el algoritmo de entrenamiento.

1. Inicializar los pesos de la red con valores aleatorios.
2. Presentar un patrón de entrada,  $X_p$ :  $X_{p1}, X_{p2}, X_{p3}, \dots, X_{pn}$ , y especificar la salida deseada que debe generar la red  $d_1, d_2, d_3, \dots, d_m$  (si la red se utiliza como un clasificador, todas las salidas deseadas serán cero, salvo una, que será la de la clase a la que pertenece el patrón de entrada).

3. Calcular la salida actual de la red, para ello se presentan las entradas a la red y se va calculando la salida que presenta cada capa hasta llegar a la capa de salida, ésta será la salida de la red  $y_1, y_2, y_3, \dots, y_m$ . Los pasos son los siguientes:

- a) Se calculan las entradas netas para las neuronas ocultas procedentes de las neuronas de entrada.

$$\text{Para una neurona } j \text{ oculta: } net_{pj}^h = \sum_{i=1}^N w_{ji}^h x_{pi} + \theta_j^h$$

En donde el índice  $h$  se refiere a magnitudes de la capa oculta (*hidden*); el subíndice  $p$ , al  $p$ -ésimo vector de entrenamiento, y  $j$  a la  $j$ -ésima neurona oculta. El término  $\theta$  puede ser opcional, pues actúa como una entrada más.

- b) Se calculan las salidas de las neuronas ocultas.

$$y_{pj} = f_j^h(net_{pj}^h)$$

- c) Se realizan los mismos cálculos para obtener las salidas de las neuronas de salida (capa u *output*).

$$net_{pk}^o = \sum_{j=1}^L w_{kj}^o y_{pj} + \theta_k^o$$

4. Calcular los términos de error para todas las neuronas. Si la neurona  $k$  es una neurona de la capa de salida, el valor de la delta es:

$$\delta_{pk}^o = (d_{pk} - y_{pk}) f_k^{o'}(net_{pk}^o)$$

Si la neurona  $j$  no es de salida, la expresión obtenida en este caso es

$$\delta_{pj}^h = f_j^{h'}(net_{pj}^h) \sum_k \delta_{pk}^o w_{kj}^o$$

Donde se observa que el error en las capas ocultas depende de todos los términos de error de la capa de salida. De aquí surge el término propagación hacia atrás.

- Actualización de pesos. Para ello, se utiliza el algoritmo recursivo, comenzando por las neuronas de salida y trabajando hacia atrás hasta llegar a la capa de entrada, ajustando los pesos de la forma siguiente:

Para los pesos de las neuronas de la capa de salida:

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \Delta w_{kj}^o(t+1); \Delta w_{kj}^o(t+1) = \alpha \delta_{pk} y_{pj}$$

Y para los pesos de las neuronas de la capa oculta:

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \Delta w_{ji}^h(t+1); \Delta w_{ji}^h(t+1) = \alpha \delta_{pj} x_{pi}$$

En ambos casos, para acelerar el proceso de aprendizaje, se puede añadir un término momento de valor  $\beta w_{kj}^o(t) - w_{kj}^o(t-1)$  en el caso de la neurona de salida, y  $\beta w_{ji}^h(t) - w_{ji}^h(t-1)$  cuando se trata de una neurona oculta.

- El proceso se repite hasta que el término de error resulta aceptablemente pequeño para cada uno de los patrones aprendidos.

$$E_p = \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2$$

## 2.2.3 Algoritmos genéticos (AG)

### 2.2.3.1 Definición

Para Herrera, Lozano y Verdegay (1994) *“los Algoritmos Genéticos son procedimientos adaptativos para la búsqueda de soluciones en espacios complejos inspirados en la evolución biológica, con patrones de operaciones basados en el principio darwiniano de reproducción y supervivencia de los individuos que mejor se adaptan al entorno en el que*

viven.” (snp) Estos algoritmos se utilizan con frecuencia para resolver problemas de búsqueda y optimización.

Asimismo estos autores consideraron que aunque existen muchas posibles variantes de los AG, los mecanismos fundamentales bajo los cuales funcionan son: operar sobre una población de individuos, que inicialmente se genera de forma aleatoria, y cambiar en cada iteración la población atendiendo a los tres siguientes pasos: (1) evaluación de los individuos de la población, (2) selección de un nuevo conjunto de individuos y reproducción sobre la base de su relativa conveniencia o adaptación, y (3) recombinación para formar una nueva población a partir de los operadores de cruce y mutación. Los individuos resultantes de estas tres operaciones forman la siguiente población, iterando este proceso hasta que el sistema cese de mejorar.

Para Ramírez (1999) los AG son algoritmos de búsqueda fundamentados en la mecánica de la selección natural y la genética de los seres vivos. Estos combinan la característica de supervivencia del más fuerte dentro de una población de individuos, utilizando un estructurado, pero aleatorio, intercambio de información para conformar un procedimiento de búsqueda que contiene algo del instinto de los seres vivos.

Por su parte Gestal (s.f.) define los Algoritmos Genéticos como métodos adaptativos, generalmente usados en problemas de búsqueda y optimización de parámetros, los cuales están basados en la reproducción sexual y en el principio de supervivencia del más apto.

Además, Gestal hace mención a la definición dada por Goldberg, la cual describe como más formal y que definió a los AG como *“algoritmos de búsqueda basados en la mecánica de selección natural y de la genética natural. Combinan la supervivencia del más apto entre estructuras de secuencias con un intercambio de información estructurado, aunque aleatorizado, para constituir así un algoritmo de búsqueda que tenga algo de las genialidades de las búsquedas humanas”* (Goldberg, 1989 citado por Gestal, s.f. p.2).

### 2.2.3.2 Componentes

Para alcanzar la solución a un problema se parte de un conjunto de individuos (que constituyen las posibles soluciones del problema), los cuales pueden representarse como un conjunto de parámetros, denominados genes, que agrupados en una lista conforman lo que se conoce como cromosomas. Normalmente estos cromosomas son representados mediante modelos binarios {0,1}, sin embargo, esto no siempre es necesario, ya que, también existen modelos para datos continuos.

Dentro de la terminología relacionada con los AG hay varios términos que es importante definir, con el fin de entender el proceso desarrollado por el algoritmo y los resultados de su aplicación. Entre estos términos, la Universidad del País Vasco señala como clave los siguientes: tamaño de la población, población inicial, función objetivo, selección, cruce y mutación.

**Tamaño de la población:** se refiere al número de soluciones que se van a estar analizando en cada proceso de escogencia; este tamaño de población no se va a variar dentro de las diferentes iteraciones del modelo, además tiene la particularidad de que si es muy pequeño puede dejar por fuera soluciones óptimas y por el contrario si es muy grande, va a generar que exista un costo mucho más alto en el proceso mismo de desarrollo del algoritmo.

**Población inicial:** se refiere a las posibles soluciones con las que se va a iniciar el proceso de trabajo, generalmente se determina aleatoriamente, de manera que no existan subjetividades en el proceso.

**Función objetivo:** es aquella función de evaluación de las alternativas de solución; esta función debe estar definida en términos de lo que se quiere optimizar, es decir, es a través de esta función que se evalúa cada una de las posibles soluciones, con el fin de encontrar la solución óptima.

**Selección:** se refiere al proceso de escogencia de los padres para la realización de los cruces; la función más utilizada es la función que se basa en la proporcionalidad con la función objetivo, es decir que aquellas soluciones que generan mejores valores de la función objetivo son las que se definen como padres para realizar las combinaciones.

Existen dos clasificaciones de los métodos de selección: los métodos dinámicos, donde las probabilidades de escoger un individuo como padre varían en cada una de las iteraciones; y los métodos estáticos, los cuales se mantienen constantes a través de las diferentes iteraciones del algoritmo.

**Cruce:** este se refiere al proceso donde se combinan dos individuos para formar dos individuos nuevos. Esto se realiza generalmente definiendo un punto de cruce de manera aleatoria y posteriormente se intercambian las partes derecha de cada solución de manera que se obtienen dos nuevos individuos, los cuales se evalúan con base en la función objetivo para determinar si se producen soluciones mejores producto de estas combinaciones.

**Mutación;** se considera un operador básico, el cual proporciona un elemento de cambio aleatorio en los individuos de la población, este cambio permite buscar más puntos en el espectro de soluciones y es especialmente efectivo cuando las soluciones evaluadas se acercan más al punto óptimo.

Existe un concepto adicional que se utiliza en algunos de los algoritmos genéticos, como lo es el de la **reducción**, este proceso consiste en seleccionar un grupo más reducido de individuos para seguir trabajando, esto en un momento definido, con el fin de trabajar con una población más pequeña. Esta reducción se realiza generalmente por un proceso elitista, de manera que aquellos individuos con mejores resultados en la función objetivo son los que prevalecen.

### **2.2.3.3 Metodología**

De acuerdo con lo definido por Zúñiga (2004), los pasos para realizar un algoritmo genético son:

#### **Población inicial**

El proceso para el desarrollo del algoritmo parte de la definición de una población inicial, la cual debe tener un tamaño tal que abarque la mayor parte del espectro de búsqueda de posibilidades, que normalmente oscila entre los 20 y 200 individuos. Para ello, primero se debe definir cómo estará representado cada individuo y luego se genera cada uno mediante un proceso aleatorio, utilizando como posibles individuos todo el espectro de solución.

#### **Evaluación de individuos**

Esta etapa permite definirle a cada individuo su fortaleza con respecto al resto de la población. Para esto se debe definir la función objetivo, que es la función que se desea optimizar; esta puede ser tanto una función de maximización de beneficios como una minimización de costos. Matemáticamente no hay requisitos mayores para esta función, lo único que se requiere es que se pueda evaluar a lo largo del espacio de búsqueda establecido para el problema.

#### **Selección**

El proceso de selección consiste en definir los individuos que mejores resultados presenten de acuerdo con el valor de la función objetivo, estos individuos son los que participan en el proceso de recombinación.

### **Recombinación o cruce**

Este paso consiste en combinar los individuos escogidos, para buscar mejorar los resultados. Este proceso se hace mediante el intercambio de genes entre los individuos para formar otros nuevos, cuya cantidad de genes se define de manera aleatoria. Para este proceso se respeta el mejor individuo de la selección anterior y se aplican los cruces sobre los siguientes mejores, de manera que se respete la ley del más fuerte, el cual sobrevive invariable. Este paso permite buscar soluciones dentro del espectro cercano a los mejores valores, con tal de explorar esta región de solución en busca de la solución óptima.

Una vez que se tienen estos nuevos individuos, se evalúan mediante la función objetivo y se organizan con respecto al valor resultante de la evaluación.

### **Mutación**

Este operador se aplica modificando aleatoriamente alguno de los genes de uno o varios individuos en particular, de manera que se tengan otros totalmente diferentes para ser evaluados. El número de individuos a los que se les aplica la mutación, así como el número de genes que se cambian y a los individuos que se les aplica, se definen de forma aleatoria y una vez más se respeta al mejor individuo, el cual permanece invariable para respetar su capacidad de ser la mejor solución encontrada hasta ese momento.

### **Evaluación de nueva generación**

Esta nueva generación de individuos obtenidos debe ser evaluada con respecto a la función objetivo, de manera que se pueda determinar su capacidad para ser una solución óptima para el problema.

Posteriormente se clasifican los individuos de acuerdo con su valor de la función objetivo y se vuelven a aplicar los pasos de cruce y mutación, hasta que los resultados obtenidos no

mejoren significativamente los resultados y se pueda definir que el individuo determinado con mejor valor de la función objetivo es el óptimo global que mejor soluciona el problema.

Es importante destacar que los individuos que se evalúen dentro del algoritmo sean soluciones factibles, para ello es necesario definir una población tal que no permita incluir individuos que no sean factibles, o por el contrario se debe definir una función objetivo de tal forma que los no factibles sean castigados en su valor, de manera que estos se vayan descartando en el proceso de iteraciones de selección de individuos.

#### **2.2.3.4 Aplicación**

Debido a la simplicidad de estos algoritmos y a la robustez de los resultados generados, este proceso puede ser aplicado fácilmente a cualquier problema de búsqueda y optimización, en cualquier campo.

De acuerdo con Zúñiga (2004), en el caso de problemas más complejos, donde se necesita realizar un análisis muy amplio de posibles soluciones o se requiere encontrar soluciones muy puntuales que generan excelentes resultados, el proceso iterativo continuo necesario para encontrar estas soluciones puede volverse lento, para ello se han tratado de ubicar nuevos métodos o formas de aplicación, tal es el caso del uso de parámetros adaptativos, algoritmos coevolutivos o híbridos, los cuales mezclan diversas técnicas de optimización; también se ha utilizado el paralelismo para solucionar los problemas de tiempos de convergencia, lo cual consiste en utilizar al mismo tiempo varios algoritmos dentro de otro general, de manera que se puedan evaluar mayor cantidad de individuos en menos tiempo y se pueda disminuir el tiempo necesario para obtener un óptimo global.

### **2.3 Marco Teórico**

Son muy pocos los autores que han combinado teoría de pronósticos, con la teoría de redes neuronales y de algoritmos genéticos en una sola obra, sobre todo por lo novedoso

de las últimas dos temáticas. Difícilmente un mismo autor ha incorporado los tres temas, pero, si hay varios que al menos han combinado dos de los tres mencionados.

Adeli y Hung (1994) en el libro *Machine Learning: Neural Networks, Genetic Algorithms, and Fuzzy Systems*, describieron el desarrollo de algoritmos de aprendizaje automático mediante la integración de las redes neuronales con otras tecnologías mejorarán sustancialmente el rendimiento.

Rojas (1996) en su libro *Neural Networks - A Systematic Introduction*, incorporó un capítulo correspondiente a los algoritmos genéticos y su relación con las redes neuronales.

Jain y Martin (1998) en el libro *Fusion of Neural Networks, Fuzzy Systems and Genetic Algorithms: Industrial Applications*, integraron las redes neuronales, la lógica difusa y los algoritmos genéticos para el diseño de sistemas complejos que consideren técnicas de computación evolutiva.

Kurkova y otros (2001) en el libro *Artificial Neural Nets and Genetic Algorithms*, presentaron puntos de vista teóricos y aplicaciones exitosas de redes neuronales combinadas con algoritmos genéticos.

La mayoría de teóricos han introducido los algoritmos genéticos en la teoría de redes neuronales, ya que, estos permiten mejorar el proceso de aprendizaje de las redes. Es decir, en vez de utilizar el algoritmo *Backpropagation*, usado tradicionalmente para el aprendizaje de la red, han utilizado un algoritmo genético para modificar los pesos.

Por otro lado, hay autores que han incorporado las redes neuronales en la teoría de pronósticos, ya que, utilizaron las redes como un método novedoso de pronósticos en diferentes campos: negocios, finanzas, etc., con el fin de considerar ciertas variables cualitativas que los métodos tradicionales omiten.

Zhang, Patuwo y otros (1998) en el artículo *Forecasting with artificial neural networks: The state of the art*, presentaron el estado del arte de la utilización de las redes neuronales para pronósticos.

Masters (1995) en su libro *Neural, Novel & Hybrid Algorithms for Time Series* abarcó tanto la aplicación práctica de métodos de pronósticos tradicionales así como técnicas innovadoras que utilizaban redes neuronales y lógica difusa.

Makridakis, Wheelwright and Hyndman (1997) en el libro *Forecasting: Methods and Applications*, presentaron una amplia gama de métodos de pronóstico, dentro de los que incluyen las redes neuronales artificiales.

Reed y Marks (1999) en el libro *Neural Smothing: Supervised Learning in Feedforward Artificial Neural Networks*, explicaron la forma de diseñar y construir redes neuronales para aplicaciones de pronósticos mediante el reconocimiento de patrones.

Zhang (2004) en el libro *Neural Networks in Business Forecasting* presentó una serie de estudios de casos para mostrar las aplicaciones innovadoras o exitosas de las redes neuronales en muchas áreas de negocios, así como métodos para mejorar la red neuronal para pronósticos.

Para el desarrollo del presente trabajo de investigación se partió de una perspectiva teórica donde se integraron las tres temáticas en cuestión. Por un lado se partió del hecho de que es posible realizar pronósticos utilizando la técnica de redes neuronales artificiales, de forma que además de variables cuantitativas consideradas en los métodos tradicionales de pronósticos, se incorporen variables cualitativas que permitan una mayor precisión de las predicciones. Por otro lado, mediante los algoritmos genéticos se entrenó la red neuronal para facilitar el aprendizaje de la misma, y por consiguiente la obtención de proyecciones con un “menor error de pronóstico” que con métodos tradicionales.

## **CAPÍTULO III. MARCO METODOLÓGICO**

El capítulo 3 del trabajo de investigación corresponde al marco metodológico. Este capítulo está estructurado en 7 apartados: tipo de estudio, sujetos y fuentes de investigación, variables, tipo de muestreo, tiempo y espacio, instrumentos, y tratamiento de la información.

### **3.1 Tipo de estudio**

De acuerdo con Grajales (1996) la investigación se puede clasificar desde tres perspectivas diferentes: según la fuente de datos, según la finalidad y según la profundidad del estudio.

Según la fuente de datos una investigación se divide en documental, de campo o experimental.

- La investigación documental se realiza apoyándose en fuentes de carácter documental, esto es, en documentos de cualquier especie.
- La investigación de campo se apoya en informaciones que provienen entre otras, de entrevistas, cuestionarios, encuestas y observaciones.
- La investigación experimental obtiene su información de la actividad intencional realizada por el investigador y que se encuentra dirigida a modificar la realidad con el propósito de crear el fenómeno mismo que se indaga, y así observarlo.

Según la finalidad las investigaciones se dividen en básica y aplicada.

- La investigación básica también recibe el nombre de investigación pura o teórica. Se caracteriza porque parte de un marco teórico y permanece en él; la finalidad radica en formular nuevas teorías o modificar las existentes, es decir, en incrementar los conocimientos científicos.

- La investigación aplicada, también llamada práctica o empírica; se caracteriza porque busca la aplicación o utilización de los conocimientos que se adquieren. Se vincula con la investigación básica, pues depende de los resultados y avances de esta última.

Según la profundidad del estudio, la investigación se divide en exploratoria, descriptiva o explicativa.

- La investigación exploratoria se realiza con el propósito de destacar los aspectos fundamentales de una problemática determinada y encontrar los procedimientos adecuados para elaborar una investigación posterior.
- La investigación descriptiva utiliza el método de análisis; mediante esta se logra caracterizar un objeto de estudio o una situación concreta, señalar sus características y propiedades. Además, puede servir de base para investigaciones que requieran un mayor nivel de profundidad.
- La investigación explicativa requiere la combinación de los métodos analítico y sintético, en conjugación con el deductivo y el inductivo. Mediante este tipo de investigación se trata de responder o dar cuenta de los porqué del objeto que se investiga.

Con base en lo anterior, la presente investigación es de tipo documental, aplicada, descriptiva y explicativa. Es documental porque para iniciar es necesario hacer una revisión detallada de libros, tesis y artículos con el fin de establecer los elementos teóricos y metodológicos relacionados con redes neuronales y algoritmos genéticos. Es aplicada ya que después de la investigación teórica se desarrollan dos aplicaciones prácticas para proyecciones sobre pacientes con cáncer hospitalizados en la CCSS. Es descriptiva ya que para diseñar los modelos de redes neuronales es necesario describir detalladamente el comportamiento del fenómeno de estudio; y finalmente es explicativa, ya que el modelo

busca explicar cómo las variables se relacionan para generar una salida determinada, es decir, tiene un componente de inferencia estadística.

Por otro lado, de acuerdo con Pita y Pértegas (2002) la investigación cuantitativa es aquella en la que se recogen y analizan datos cuantitativos sobre variables; trata de determinar la fuerza de asociación o correlación entre variables, la generalización y objetivación de los resultados a través de una muestra para hacer inferencia a una población de la cual toda muestra procede. Por otra parte, según los mismos autores, la investigación cualitativa hace la descripción de las cualidades de un fenómeno que son estudiados mediante técnicas como la observación participante y las entrevistas no estructuradas. Se trata de obtener un entendimiento lo más profundo posible.

Este proyecto corresponde a una investigación cuantitativa, ya que, el objetivo fue el desarrollo de un modelo de pronósticos mediante redes neuronales, lo cual implica la relación de diferentes variables para obtener un resultado numérico de la proyección; además se optimizaron las redes con algoritmos genéticos y se compararon con otros modelos tradicionales analizando los errores de pronósticos, medido todo en términos cuantitativos.

A pesar que la investigación es cuantitativa tiene un componente cualitativo, utilizado para identificar las variables que podrían afectar el modelo de proyección, mediante las entrevistas no estructuradas con expertos en el tema de redes neuronales, así como en el tema de la salud y de la incidencia en esta de variables socioeconómicas que se podrían incluir en el análisis.

## **3.2 Sujetos y fuentes**

### **3.2.1 Sujetos**

Los sujetos de esta investigación son los pacientes con cáncer internados en los diferentes hospitales de la Caja Costarricense de Seguro Social. El objetivo es que el modelo de redes neuronales permita hacer proyecciones sobre este tipo de pacientes, como la cantidad de casos nuevos hospitalizados por mes o la estancia de hospitalización de los mismos, razón por la cual fue necesario analizar las variables internas y externas que afectan a este tipo de pacientes.

Se tomaron en cuenta pacientes de todas las edades, de ambos sexos, cuyo diagnóstico principal para la hospitalización fuera cualquiera de los posibles tumores malignos que existen, de acuerdo con la Clasificación Internacional de Enfermedades, en su décima edición (CIE 10).

### **3.2.2 Fuentes**

Para el diseño teórico y aplicación metodológica del modelo de redes neuronales la fuente de información principal fue la recopilación bibliográfica, por medio de la consulta de libros, tesis, revistas, artículos e Internet. Adicionalmente se realizó una entrevista no estructurada a expertos en el área técnica y en el tema de análisis.

Para el diseño funcional del modelo, la fuente de información principal correspondió a la base de datos de egresos hospitalarios de la Caja Costarricense de Seguro Social, en la cual se dispone de información de los pacientes hospitalizados con cáncer atendidos en dicha Institución. El acceso a esta base de datos se realizó por medio del Área de Estadística en Salud de la CCSS, la cual facilitó los datos de manera digital.

Esta base de datos está constituida por 98 variables y tiene información sobre 119132 egresos de pacientes con cáncer, para el período comprendido entre los años 1997 y 2010.

Adicionalmente, para el estudio de las variables macroeconómicas y sociales que afectan el fenómeno en estudio se consultaran fuentes como el Ministerio de Planificación (MIDEPLAN), el Instituto Nacional de Estadística y Censos (INEC) y el Programa de Naciones Unidas para el Desarrollo (PNUD). Los datos fueron obtenidos por medio de Internet, en las páginas Web de estas instituciones.

Finalmente, se realizó una triangulación de resultados: los que genera el modelo con base en los datos que lo alimentan, los que genera un programa computacional especializado y los datos reales de pacientes con cáncer atendidos en la CCSS.

### **3.3 Tiempo y espacio del proyecto**

El proyecto está definido en el contexto del año 2011, en donde se utilizó como base de análisis y complemento de la misma, las investigaciones y procedimientos que se encontraron debidamente documentados y publicados a ese momento, especialmente sobre los temas involucrados en el trabajo que son de más reciente desarrollo, como las redes neuronales, los algoritmos genéticos y la interacción que se pueda generar entre ellos.

Por otro lado, para la construcción y entrenamiento del modelo, se utilizaron datos sobre los pacientes con cáncer hospitalizados en la CCSS en los últimos 14 años (1997-2010) así como los datos sobre las variables macroeconómicas y sociales del país de los últimos 17 años (1994-2010). Se analizaron estos períodos, ya que, sobre estos son las que las instituciones consultadas tienen datos disponibles.

Para analizar la efectividad del modelo, es decir para realizar la comparación con los métodos tradicionales, se utilizaron los datos del año 2010, por lo que este período no se incluyó en la construcción ni entrenamiento del modelo, en esta fase.

En cuanto al espacio de la investigación, el modelo de redes neuronales desarrollado es para realizar proyecciones sobre pacientes con cáncer atendidos en los servicios de hospitalización en la CCSS, específicamente sobre la cantidad de casos nuevos de pacientes hospitalizados en dicha institución por mes, y sobre la cantidad de días de hospitalización de estos pacientes. Sin embargo, la metodología aplicada, permite que se creen las bases para desarrollar aplicaciones posteriores sobre otras características de los pacientes con cáncer, o sobre otro tipo de patologías diferentes al cáncer.

Incluso, con algunas consideraciones adicionales, el modelo podría aplicarse en otras organizaciones, nacionales o internacionales, de diferentes tipos de industria, para realizar proyecciones sobre demanda.

### **3.4 Variables**

Dentro del proyecto hay un apartado correspondiente al análisis de las variables que afectan el fenómeno de estudio, pero, para una mejor comprensión del mismo y de todo el proceso de construcción, entrenamiento y evaluación del modelo de redes neuronales en el siguiente cuadro se presenta la conceptualización, operacionalización e instrumentalización de las mismas.

**Cuadro 2: Variables de la investigación**

Variable	Definición conceptual	Definición operacional	Definición instrumental
<b>Relacionadas con los pacientes con cáncer</b>			
Centro de Salud	Se refiere al hospital de la C.C.S.S por el cual ingresó el paciente con cáncer.	Se seleccionaron los 12 centros de salud que atendieron la mayor la cantidad de casos en el período 1997 a 2010 (91,1% del total); los casos de otros centros no fueron considerados, con el fin disminuir la variabilidad.	Esta variable es obtenida de la base de datos de egresos hospitalarios, sin embargo, aunque su codificación original incluye 4 dígitos, para efectos del desarrollo de la investigación se recodificó con números del 1 al 12 (correspondientes a los centros seleccionados).
Sexo	Se refiere al sexo de cada paciente con cáncer.	En el estudio se contemplan tanto hombres como mujeres. La simbología para el hombre es la letra "M" mayúscula (masculino) y para la mujer la letra "F" mayúscula (femenino).	Esta variable se obtiene de la base de datos de egresos hospitalarios, y se utiliza la siguiente codificación: para el hombre el número 1 y para la mujer el número 2.
Edad del paciente	Se refiere a los años de vida del paciente con cáncer en el momento del egreso del hospital de la C.C.S.S.	Se consideran los pacientes de todas las edades que presentaron un egreso de los hospitales de la CCSS en el período 1997-2010. Sin embargo, para disminuir la variabilidad de la red se agruparon los pacientes en 20 categorías, con rangos de 5 años cada una.	Esta variable es obtenida de la base de datos de egresos hospitalarios, sin embargo, aunque su codificación original es de 5 dígitos (horas, días, meses, años, cientos de años), para efectos del proyecto se recalculó a años corrientes y posteriormente se recodificó con números del 1 al 20 (correspondientes a las 20 categorías o rangos de edades establecidos).
País de nacimiento	Se refiere al país en el que nació el paciente con cáncer que ingresa a la CCSS.	Para efectos de la presente investigación se consideran todos los pacientes independientemente de su país de nacimiento-	Esta variable es obtenida de la base de datos de egresos hospitalarios, y su codificación responde a la identificación internacional de cada país.
Residencia	Se refiere al lugar en cual habita el paciente con cáncer al ingresar en uno de los centros de salud de la C.C.S.S.	Aunque la residencia incluye tanto la provincia como el cantón y el distrito, con el fin de disminuir la variabilidad de los datos, en este proyecto se considera únicamente el cantón.	Esta variable es obtenida de la base de datos de egresos hospitalarios, sin embargo, aunque su codificación original incluye 5 dígitos, para este proyecto se recodificó con números del 1 al 81 (número total de cantones que tiene nuestro país).

Variable	Definición conceptual	Definición operacional	Definición instrumental
Área de ingreso	Se refiere al área de atención en el centro de salud por la cual ingresa el paciente con cáncer.	Cuando un paciente con cáncer ingresa a uno de los centros de salud de la C.C.S.S, puede hacerlo por tres áreas posibles: consulta externa, urgencias y nacimientos.	Esta variable es obtenida de la base de datos de egresos hospitalarios y tiene una codificación con números del 1 al 3 (1 para consulta externa, 2 para urgencias y 3 para nacimientos)
Servicio de ingreso	Se refiere al servicio del centro de salud en el cual el paciente con cáncer fue atendido después de haber ingresado.	Hay 9 posibles servicios por los cuales puede ingresar un paciente con cáncer, para efectos de esta investigación e consideran los pacientes ingresados por cualquiera de estos 9 servicios.	Esta variable es obtenida de la base de datos de egresos hospitalarios y se utiliza la siguiente codificación: Medicina (1), Cirugía (2), Gineco-obstetricia (3), Pediatría (4), Cirugía Mayor Ambulatoria (6), Psiquiatría (7), Cuidados intensivos en medicina (8), Cuidados intensivos quirúrgicos (9), Cirugía Vespertina (11).
Especialidad de ingreso	Se refiere a la especialidad de las ciencias de la salud que atendió inicialmente al paciente con cáncer que ingresó a la CCSS.	Hasta la fecha se tienen identificadas 71 especialidades de ingreso a los centros de salud de la CCSS. En este proyecto se consideran todas las posibilidades.	Esta variable es obtenida de la base de datos de egresos hospitalarios y se utiliza una codificación del número 1 al 91, ya que, hay 20 especialidades que dejaron de considerarse como tales y se excluyeron.
Ingresos previos	Se refiere al número de ocasiones en que un paciente con cáncer ha ingresado a los centros de salud de la C.C.S.S antes del ingreso actual	Se contabilizan todos los ingresos del paciente en los 14 años que incluye el estudio (1997-2010)	La codificación de esta variable es numérica, iniciando en 1 hasta la cantidad final de ingresos previos por paciente. La misma es obtenida directamente de la base de datos de egresos hospitalarios.
Diagnóstico principal	Se refiere al tipo de cáncer principal sobre el cual fue diagnosticado el paciente.	Existen más de 450 clasificaciones para el cáncer. Para disminuir la variabilidad se seleccionaron los 20 diagnósticos que se presentaron en la mayor la cantidad de casos en el período 1997 a 2010 (66% del total); los casos de otros tipos de cáncer no fueron considerados	Esta variable es obtenida de la base de datos de egresos hospitalarios, sin embargo, aunque su codificación original es alfanumérica (una letra y tres números), para este proyecto se recodificó con números del 1 al 20 (correspondientes a los tipos de diagnóstico seleccionados).

Variable	Definición conceptual	Definición operacional	Definición instrumental
Estancia preoperatoria	Se refiere a la cantidad de días que pasan desde que el paciente ingresa al centro de salud hasta que es intervenido quirúrgicamente debido a que se le diagnosticó cáncer	La estancia preoperatoria se cuenta a partir del día 0 (día de ingreso) hasta el día previo a la operación. Para efectos de la presente investigación se hizo una agrupación de los días en 8 categorías: 0 a 5, 5 a 10, 10 a 15, 15 a 30, 30 a 45, 45 a 60, 60-90, más de 90.	Esta variable es obtenida de la base de datos de egresos hospitalarios. Esta variable fue recodificada con números del 1 al 8 (correspondientes a cada categoría o rangos de días).
Procedimiento realizado	Se refiere al tipo de procedimiento (médico o quirúrgico) que se le realiza al paciente posteriormente a que se le diagnostica con algún tumor maligno.	Hasta la fecha se tienen identificadas 3499 tipos de procedimientos. En este proyecto se consideran todas las posibilidades.	Esta variable es obtenida de la base de datos de egresos hospitalarios y tiene una codificación del número 1 a al 9999.
Condición de salida	Se refiere a la causa del egreso del paciente con cáncer del centro de salud de la CCSS.	Hay 7 posibilidades de egreso: indicación médica, exigida, fuga, transferido, fallecido con autopsia, fallecido sin autopsia y fallecido traslado forense. En este proyecto se consideran las 7 posibilidades.	Esta variable es obtenida de la base de datos de egresos hospitalarios y está codificada con números del 1 al 7 (correspondientes a cada posibilidad en el orden indicado en la definición operacional).
Cause básica de muerte	Se refiere al tipo de cáncer que causó la muerte de los pacientes con cáncer que egresaron fallecidos de los centros de salud de la CCSS.	Al igual que en el caso de los diagnósticos, existen 450 posibles causas de muerte. En esta investigación se consideran todas.	Esta variable es obtenida de la base de datos de egresos hospitalarios, y su codificación es alfanumérica (una letra y tres números).
Estancia total	Corresponde al número de días de hospitalización de los pacientes con cáncer atendidos en la CCSS.	Esta variable se calcula restando la fecha de egreso del paciente a la fecha de ingreso. Es decir se cuentan los días desde que el paciente ingreso hasta que egresó. Para efectos del proyecto se crearon 20 categorías con rangos de 5 días cada una, con el fin de disminuir la variabilidad de la red.	La fuente de información para el cálculo de esta variable es la base de datos de egresos hospitalarios de la C.C.S.S. Esta variable tiene una codificación con números del 1 al 20 (correspondiente a las 20 categorías definidas).

Variable	Definición conceptual	Definición operacional	Definición instrumental
Casos nuevos	Se refiere al número de casos nuevos de pacientes con cáncer hospitalizados cada mes en la CCSS.	Se calcula sumando la cantidad de pacientes con cáncer que ingresaron por primera vez a un centro de salud de la C.C.S.S en cada mes.	La fuente de información para el cálculo de esta variable es la base de datos de egresos hospitalarios de la C.C.S.S.
<b>Relacionadas con los variables macroeconómicas y sociales que afectan el fenómeno de estudio</b>			
Esperanza de vida al nacer.	Corresponde a los años esperados de vida de una persona nacida en Costa Rica.	Se calcula como el promedio ponderado de los años que vive una persona, por lo que se expresa en términos de números continuos, ya que, puede incluir decimales.	Los datos que se utilizan son obtenidos del MIDEPLAN y el PNUD
Tasa general de mortalidad.	Corresponde a la tasa de defunciones en Costa Rica	Su cálculo corresponde al número de defunciones por cada mil habitantes. Esta variable se expresa en términos de números continuos, ya que, puede incluir decimales.	Los datos utilizados en el análisis son obtenidos del MIDEPLAN.
Cantidad de Población.	Corresponde al total de personas que habitan el territorio nacional	Se establece con base en censo nacional que se realiza periódicamente, en el cual se cuentan cada una de las personas que habitan en el país. Se expresa en números enteros.	Los datos utilizados provienen del INEC
Índice de Desarrollo Humano.	Es un indicador social estadístico compuesto por tres parámetros: vida larga y saludable, educación y nivel de vida digno.	Este índice es calculado por el PNUD mediante la aplicación de diferentes instrumentos que miden el nivel de desarrollo de los países de todo el mundo. Esta variable se expresa con números continuos entre 0 y 1.	Los datos utilizados son obtenidos del MIDEPLAN, que hace referencia a los informes del PNUD.
Inversión Social en Salud.	Corresponde a la inversión del país en el tema de salud	Esta variable está definida como el porcentaje del Producto Interno Bruto (PIB) que el país invierte en programas de salud.	Los datos utilizados son obtenidos del MIDEPLAN
Tasa de escolaridad y nivel educativo.	Corresponde al nivel educativo de la población residente en el país	Esta variable se mide como la tasa bruta de escolaridad para educación diversificada, es decir el porcentaje de la población que tiene concluida la educación diversificada (colegio)	Los datos son obtenidos del MIDEPLAN

Variable	Definición conceptual	Definición operacional	Definición instrumental
Cobertura Régimen de Enfermedad y Maternidad.	Se refiere a la cobertura que tiene en el país el régimen de enfermedad y maternidad de la CCSS.	Corresponde al porcentaje de la población del país que está cubierta por dicho régimen.	Los datos son obtenidos del MIDEPLAN.
Desigualdad en la distribución del ingreso.	Corresponde a la distribución de los recursos generados en el país entre la población residente	Está medido por el coeficiente de concentración de Gini, el cual tiene un valor entre 0 y 1, donde a medida que el valor se acerca a 1, la concentración de los ingresos se hace mayor	Los datos fueron obtenidos también del MIDEPLAN
Tasa de desempleo	Corresponde al porcentaje de la población económicamente activa que no tiene trabajo.	Se calcula como el número de habitantes de la población económicamente activa que están desocupados dividido por el total de la población económicamente activa	Los datos son obtenidos directamente del MIDEPLAN.
Nivel de pobreza	Corresponde al porcentaje de la población en la que sus ingresos económicos no alcanzan para suplir las necesidades más básicas de vida.	Se mide como el porcentaje de personas que poseen un ingreso económico por debajo del límite de pobreza definido por el Banco Mundial, que es aproximadamente US\$2 diarios.	Los datos son obtenidos de varias fuentes, tales como el MIDEPLAN; PNUD, el INEC y el Banco Mundial.

Fuente: Los autores

**Cuadro 3: Variables internas del modelo**

Variable	Definición conceptual	Definición operacional	Definición instrumental
<b>Pronósticos</b>			
Error de pronóstico	Diferencia entre el valor pronosticado y el valor real de una variable dada, para un período específico.	Se obtiene al calcular la diferencia entre el valor real y el pronosticado.	El valor es obtenido a través de la aplicación automatizada.
<b>Redes Neuronales</b>			
Estado de activación	Es la variable que define la condición de una neurona en un momento dado.	Se define los estados que van a tener cada neurona, para saber si interviene o no en el modelo.	Se define en el modelo.
Peso de neuronas ( $w_{ji}$ )	Es el efecto que tiene una neurona en su influencia hacia otra neurona.	Para obtener este peso se utilizan los algoritmos genéticos, para lo cual se comparan lo resultado de acuerdo con el error de	Se define en el modelo como resultado del algoritmo genético.

Variable	Definición conceptual	Definición operacional	Definición instrumental
		pronóstico.	
<b>Algoritmos genéticos</b>			
Tamaño de población	Se refiere al número de soluciones que se van a estar analizando en cada proceso de escogencia.	Se define un tamaño que permita abarcar una zona importante del área de los datos y no genere mucha carga para la aplicación automatizada.	Se define en el modelo.
Población inicial	Se refiere a las posibles soluciones sobre las que se va a iniciar el proceso de trabajo.	Se define un tamaño que permita abarcar una zona importante del área de los datos desde un inicio.	Se define en el modelo.
Iteraciones	Se refiere al número de procesos que se van a aplicar para obtener los resultados definitivos.	Se define un tamaño manejable que permita obtener un resultado óptimo o aceptable en un tiempo adecuado.	Se define en el modelo.

Fuente: Los autores

### 3.5 Tipo de muestreo

Los datos para el desarrollo de los modelos de redes neuronales se seleccionaron mediante un muestreo intencional y no probabilístico, de acuerdo a la disponibilidad de los mismos en la base de datos de egresos hospitalarios de la CCSS, y a la disponibilidad de datos sobre las variables macroeconómicas y sociales que afectan el fenómeno de estudio.

Para la construcción del modelo de redes neuronales de casos nuevos se utilizaron todos los datos mensuales disponibles sobre la cantidad de pacientes nuevos con cáncer hospitalizados la CCSS en el período 1997-2010, así como las variables macroeconómicas indicadas para este mismo período de estudio. En total se consideran 168 datos.

Para la fase de entrenamiento se consideraron el 71% de los datos (los primeros 120) y para la fase de validación o prueba se utilizaron el 21% (siguientes 36 datos). Para evaluar la efectividad del modelo y compararlo con otros métodos tradicionales se usaron los últimos 12 datos, es decir el último año disponible.

En el caso del modelo de redes neuronales para estimación de la estancia hospitalaria, se trabajó con los datos disponibles sobre los pacientes con cáncer egresados únicamente en el año 2010, ya que, a diferencia de la otra red, la variable tiempo no tiene incidencia en el fenómeno de estudio. En este caso podrían afectar otras variables relacionadas como proceso, gestión interna o las características propias del paciente, por lo que al usar datos del 2010 se busca que el modelo represente la realidad actual de los hospitales y de la salud de la población en general.

Considerando la capacidad del programa informático utilizado, se seleccionaron 1000 casos para el modelado de esta red. El 85 % (850) de estos datos se utilizó para la construcción y evaluación del modelo y el último 15% (150 datos) para comparar las proyecciones del mismo.

### **3.6 Descripción de los instrumentos**

Para el desarrollo del modelo se contó con algunos instrumentos los cuales ayudaron a obtener y ordenar la información necesaria.

#### **3.6.1 Entrevista no estructurada**

Se realiza sin prever las preguntas, es decir, que ni las preguntas ni las respuestas están predeterminadas completamente. La finalidad de esta entrevista fue obtener el criterio de experto respecto a las variables e interrelaciones que se deben incorporar en el modelo para generar mejores resultados en las proyecciones sobre pacientes con cáncer.

#### **3.6.2 Listado o base de datos históricos**

Con base en la información sobre las variables relacionadas con los pacientes con cáncer atendidos en la CCSS, obtenida de la base de datos de egresos hospitalarios, y en los datos sobre las variables macroeconómicas y sociales del país que afectan el fenómeno de estudio, obtenidos de las diferentes instituciones mencionadas anteriormente, se hizo

necesario realizar una selección y ordenamiento de las variables en orden cronológico de modo que se facilitara su incorporación en los modelos de redes neuronales. Estos datos primero se revisaron y se depuraron, además se tabularon y se ordenaron en tablas que para incorporarlos en la herramienta automatizada.

### **3.6.3 Aplicación automatizada del modelo**

Como parte de esta investigación fue necesario desarrollar una aplicación automatizada del modelo de redes neuronales en un programa informático, la cual permite manejar los datos a gran escala para el entrenamiento de la red y facilita el proceso de optimización del modelo mediante el algoritmo genético.

Esta aplicación además permitió utilizar diferentes variables y procedimientos para mejorar los resultados y facilitar la generación de proyecciones sobre pacientes con cáncer, el cálculo de los errores de pronóstico y su posterior comparación con los modelos tradicionales.

## **3.7 Tratamiento de la información**

Una vez diseñado el modelo, así como su herramienta automatizada, se procedió al entrenamiento de la red alimentándola con los datos disponibles sobre los pacientes con cáncer atendidos en la CCSS, además de variables macroeconómicas y sociales del país.

El modelo se entrenó haciendo corridas del mismo con el mayor número posible de datos, y ajustando los pesos de modo que el resultado obtenido se igualara al valor real.

Posteriormente el modelo de redes neuronales se optimizó mediante un algoritmo genético, el cual permitió establecer el conjunto de pesos iniciales para la red neuronal, de manera que se minimizara el error a un valor óptimo.

Lo siguiente fue hacer las predicciones con base en los datos disponibles. Para esto se introdujeron los datos en la herramienta automatizada y se calcularon las proyecciones del siguiente año.

Finalmente, utilizando un programa informático confiable para pronósticos se realizó las mismas proyecciones mediante distintos tipos de pronósticos. Se debió calcular los errores de pronóstico obtenidos en cada caso y elaborar los gráficos con las proyecciones, de manera que se pudiera comparar los resultados y determinar cual método era más efectivo.

## **CAPÍTULO IV. DESARROLLO DE LA INVESTIGACIÓN**

### **4.1 Descripción del fenómeno en estudio**

El objetivo principal de la presente investigación es ampliar el conocimiento teórico en relación con la utilización de las redes neuronales y los algoritmos genéticos en el desarrollo de modelos de proyección.

Cómo fenómeno de estudio principal se estableció trabajar con los casos nuevos de pacientes con cáncer hospitalizados en la Caja Costarricense de Seguro Social (CCSS). Este tema fue definido debido a la relevancia que tiene para la institución mencionada y para el país, generar información valiosa sobre el cáncer, que permita una mejor planificación de los servicios de salud para un abordaje oportuno de la enfermedad.

Adicionalmente, se estudió la estancia de los pacientes con cáncer en los hospitales de la C.C.S.S, ya que, además de contribuir con la planificación de los servicios que mantiene la Institución, permite mostrar como mediante esta técnica reciente como lo son las redes neuronales, pueden modelarse diferentes fenómenos para generar información clave para la toma de decisiones.

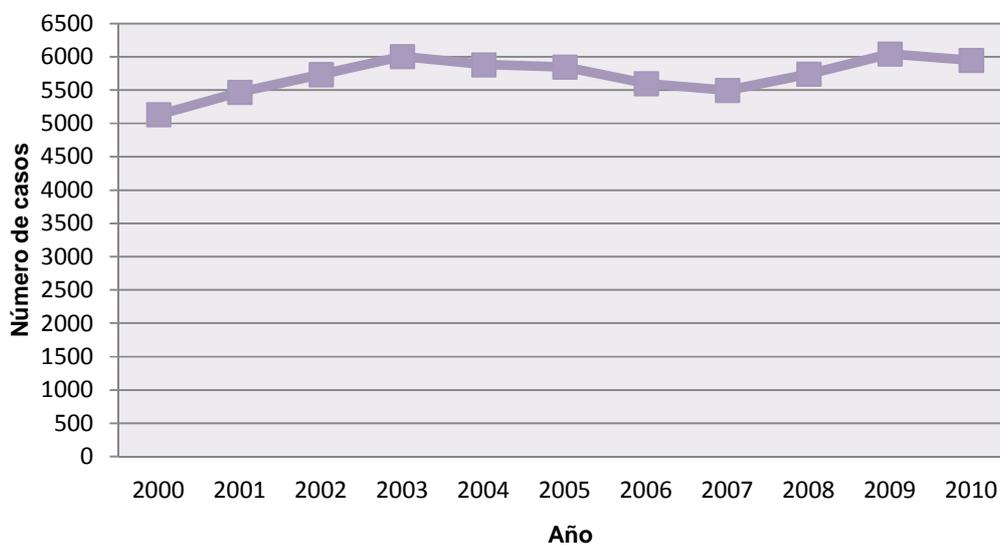
A diferencia del fenómeno de estudio principal, en que el comportamiento de casos nuevos de pacientes hospitalizados con cáncer se mide por mes, para el caso de la estancia la red debía considerar los datos de cada paciente individual. Esto, además de mostrar la versatilidad de las redes neuronales, muestra como mediante esta técnica es posible crear aplicaciones para el manejo de grandes volúmenes de datos entre los cuales las relaciones entre las variables son casi imperceptibles al ojo humano.

A continuación se describe el comportamiento observado de las variables principales relacionadas con el fenómeno de estudio para el período comprendido entre los años 2000 y 2010.

#### 4.1.1 Evolución del número de casos nuevos de pacientes hospitalizados con cáncer

Desde el año 2000 hasta el 2010 la cantidad de casos nuevos de pacientes con cáncer hospitalizados en la CCSS ha oscilado entre 400 y 500 mensuales, lo que es equivalente a entre 5000 y 6000 casos anuales. Esto se observa en el siguiente gráfico.

**Gráfico 1 : Evolución del número de casos nuevos de pacientes con cáncer hospitalizados en la CCSS entre los años 2000 y 2010.**



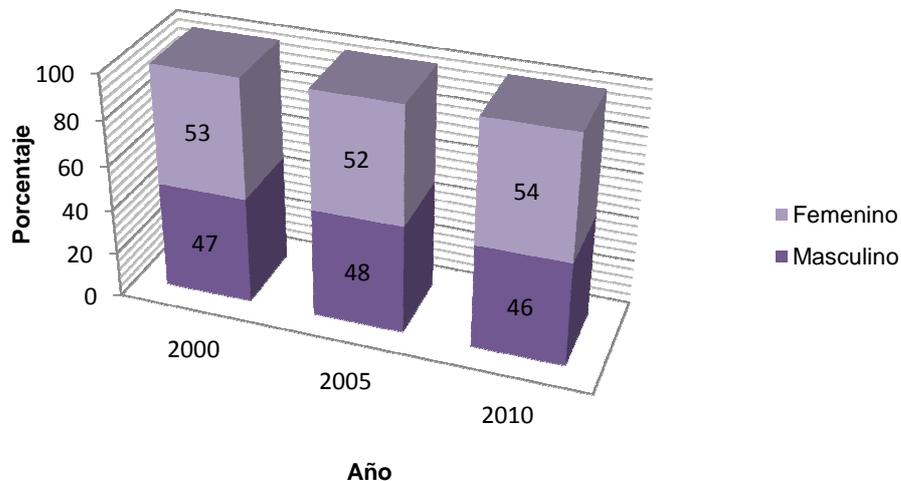
Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

Como se observa en el gráfico anterior, del año 2000 al 2003 se presentó un incremento gradual en la cantidad de casos hasta llegar a los 6006 en el 2003. A partir de este año se presentó una tendencia a disminuir el número de pacientes nuevos con cáncer hospitalizados en la CCSS, hasta llegar a 5499 en el 2007. En el 2009 se atendió la mayor cantidad de casos considerando todo el período de estudio, se atendieron 6045 casos; finalmente, en el 2010 esta cantidad disminuyó ligeramente hasta los 5946 casos.

#### 4.1.2 Sexo

En el siguiente gráfico se presenta la distribución, según sexo, de la cantidad de pacientes nuevos con cáncer hospitalizados en la CCSS, considerando tres años como punto de comparación: 2000, 2005 y 2010.

**Gráfico 2 : Distribución porcentual de la cantidad de casos nuevos de pacientes con cáncer hospitalizados en la CCSS, según sexo para el período 2000-2010.**



Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

Como se observa en el gráfico 2, en los tres períodos analizados se atendieron más mujeres con cáncer que hombres, con una distribución porcentual muy similar. En el año 2000 se atendieron 53% de casos en que los pacientes fueron mujeres y 47% hombres; en el 2005 se atendieron 52% de mujeres y 48% hombres, y en el 2010 el 54% de los casos fueron de mujeres y 46% de hombres.

#### 4.1.3 País de nacimiento

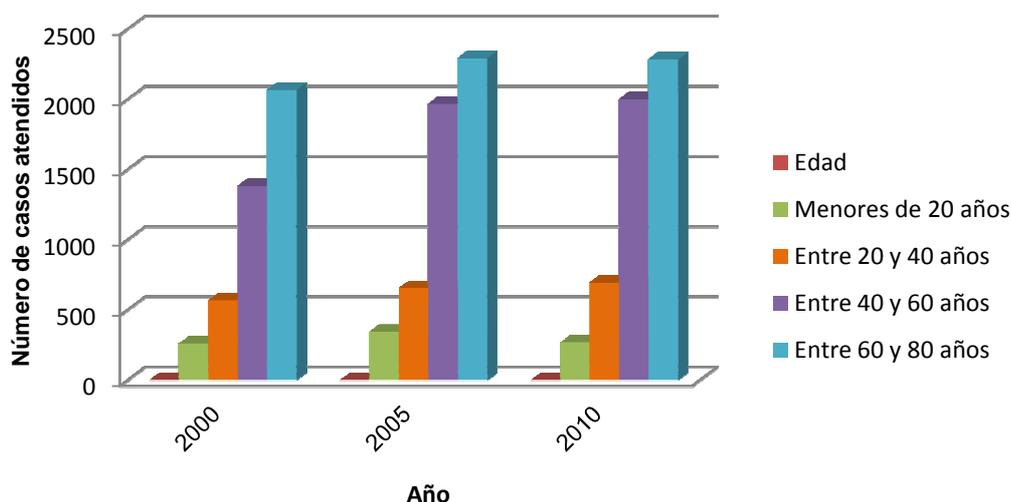
En los años 2005 y 2010 del total de pacientes nuevos con cáncer hospitalizados en la CCSS, el 93% nacieron en Costa Rica (5453 y 5522 respectivamente), 4,8% en Nicaragua (281 y 283 respectivamente) y el restante 2% en países como Estados Unidos, Colombia, El

Salvador, Panamá, entre otros. Del año 2000 no se tienen datos, ya que, en ese período no se registraba en la base de datos el país de nacimiento de los pacientes.

#### 4.1.4 Edad

Como se observa en el siguiente gráfico, más del 70% de los pacientes con cáncer hospitalizados en la CCSS tienen edades entre los 40 y los 80 años. Cerca de 5 % de los pacientes son menores de 20 años, 12% tienen entre 20 y 40 años y 12% tienen más de 80 años. Esta distribución porcentual es similar en los tres años de estudio.

**Gráfico 3 : Distribución del número de casos nuevos de pacientes con cáncer hospitalizados en la CCSS, según edad para el período 2000-2010.**



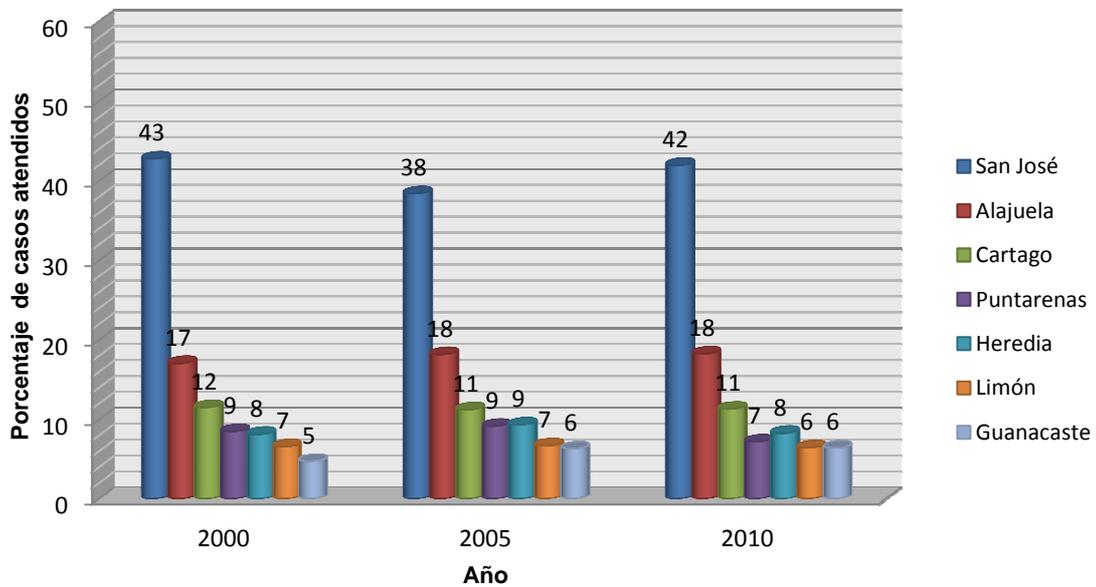
Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

En el año 2000, 42% de los pacientes nuevos hospitalizados tenían edades entre los 60 y los 80 años (2066); en el 2005 fue 39% de los pacientes los que se encontraban en ese rango de edades (2293) y en el año 2010 fue 38% (2285). Para el rango de edad entre 40 y 60 años, la cantidad de casos en el 2000 fue de 1382, en el 2005 fue de 1967 y en el 2010 fueron 2001 pacientes.

#### 4.1.5 Residencia

La distribución según residencia de la cantidad de pacientes nuevos con cáncer hospitalizados en la CCSS, está directamente relacionada con distribución de la población del país por provincia. Como se observa en el siguiente gráfico, para los años 2000, 2005 y 2010 la mayor cantidad de pacientes residían en San José (43%, 38% y 42% respectivamente). Luego le sigue Alajuela con 17%, 18% y 18% del total de casos en los tres años indicados.

**Gráfico 4 : Distribución porcentual de la cantidad de casos nuevos de pacientes con cáncer hospitalizados en la CCSS, según provincia de residencia para el período 2000-2010.**



Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

Del gráfico también se observa como Cartago le seguía a San José y Alajuela como la tercera provincia en la que residían la mayoría de los pacientes nuevos con cáncer hospitalizados de la CCSS, con un total del 12%, 11% y 11% respectivamente.

En cuanto a la residencia por cantón, como se observa en el siguiente cuadro, en el año 2010 once cantones aportaron el 50% del total de pacientes nuevos con cáncer

hospitalizados en la CCSS. Entre estos ocho, hay cinco cabeceras de provincia (San José, Alajuela, Cartago, Heredia y Liberia) y los restantes seis son ciudades que al igual que las cabeceras de provincia tienen una gran concentración poblacional: Desamparados, Goicoechea, Pérez Zeledón, San Carlos, Tibás y Montes de Oca. Es importante indicar que de los once cantones, seis pertenecen a San José. La distribución de los pacientes según cantón de residencia es muy similar para los otros años de estudio.

**Cuadro 4: Principales cantones de residencia de los pacientes nuevos con cáncer hospitalizados en la CCSS en el año 2010.**

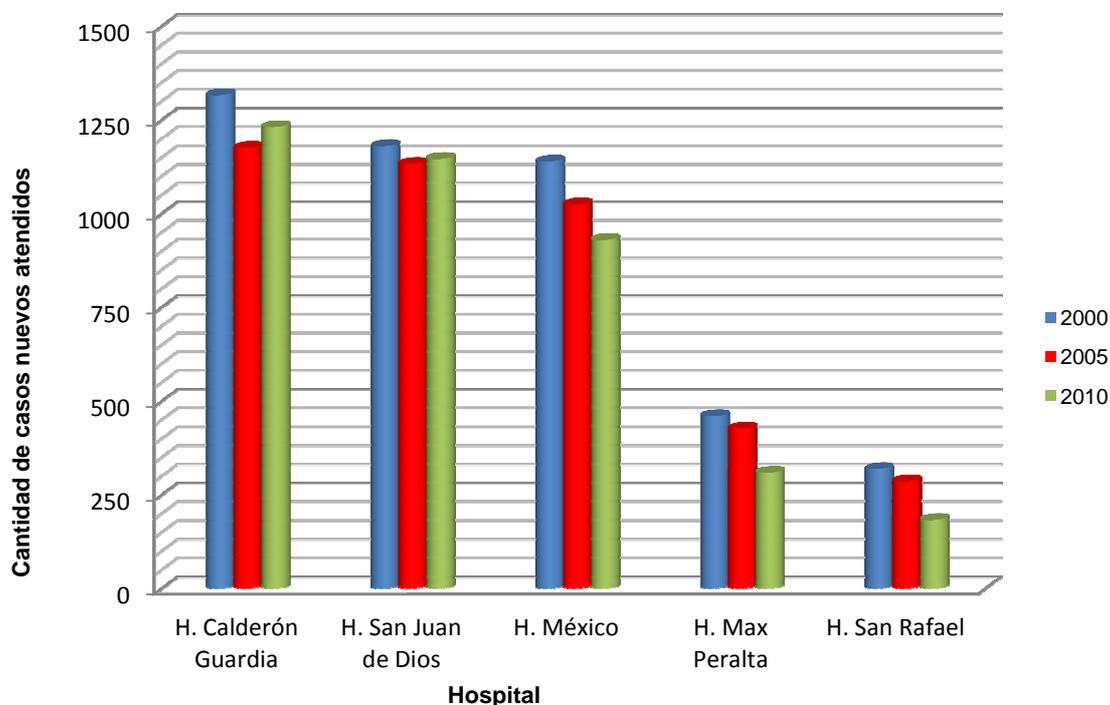
Residencia	Frecuencia	Porcentaje	Porcentaje acumulado
San José	594	10%	10%
Alajuela	430	8%	18%
Desamparados	315	6%	23%
Goicoechea	239	4%	28%
Cartago	227	4%	32%
Pérez Zeledón	202	4%	35%
Liberia	197	3%	39%
San Carlos	177	3%	42%
Heredia	141	2%	44%
Tibás	136	2%	47%
Montes de Oca	124	2%	49%
Otros	3164	51%	100%
<b>Total</b>	<b>5946</b>	<b>100%</b>	

Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

#### 4.1.6 Centros de salud de ingreso

La mayor cantidad de casos nuevos de pacientes con cáncer son internados en los hospitales nacionales generales de la CCSS. El hospital que más casos atiende es el Calderón Guardia, le sigue el San Juan de Dios y posteriormente el México. Además de los tres anteriores, los hospitales Max Peralta y San Rafael de Alajuela atienden un número considerable de casos. Este comportamiento se repite en el año 2000, 2005 y 2010, tal como se observa en el siguiente gráfico.

**Gráfico 5 : Centros de salud de la CCSS en los que son hospitalizados la mayor cantidad de pacientes nuevos con cáncer, en el período 2000-2010.**



Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

En el año 2000, los cinco centros mencionaron atendieron 4414 casos (74% del total). El Hospital Calderón Guardia atendió 1316 casos, el San Juan de Dios 1180 casos, el México 1139 casos, el Max Peralta 460 casos y el San Rafael de Alajuela 319.

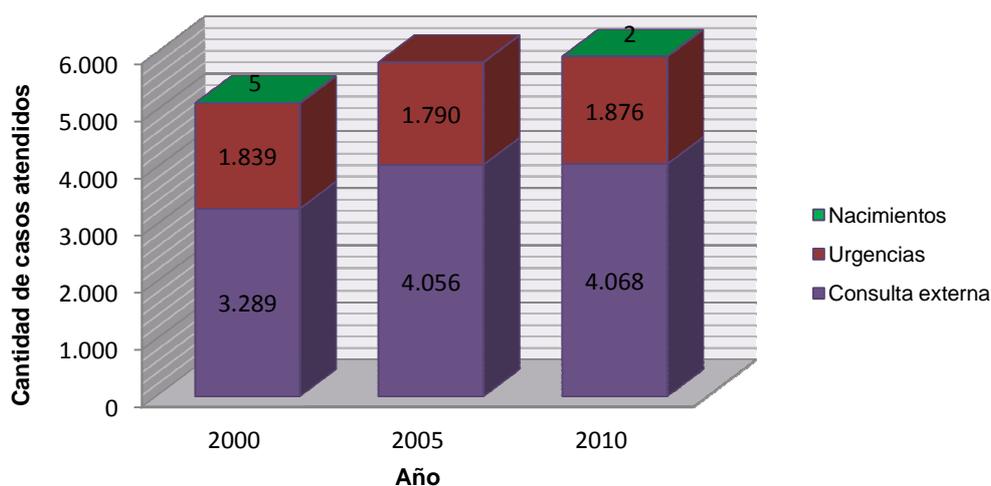
En el año 2005 el Hospital Calderón Guardia atendió 1176 casos, el San Juan de Dios 1133 casos, el México 1025 casos, el Max Peralta 427 casos y el San Rafael de Alajuela 286 casos. Esto equivale a un 69% del total de casos.

Finalmente en el año 2010 los centros de salud anteriormente citados atendieron 3799 casos (74% del total) distribuidos de la siguiente forma: 1231 en el Calderón Guardia, 1146 en el San Juan de Dios, 930 en el México, 309 en el Max Peralta y 183 en el San Rafael de Alajuela.

#### 4.1.7 Área de ingreso

En el siguiente gráfico se presenta la distribución de los casos de pacientes nuevos con cáncer hospitalizados en la CCSS, según el área de ingreso al centro de salud.

**Gráfico 6 : Distribución del número de casos de pacientes nuevos con cáncer hospitalizados en la CCSS, según área de ingreso para el período 2000-2010.**



Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

Como se observa en el gráfico anterior, la mayor cantidad de pacientes hospitalizados ingresan por el área de consulta externa. En el año 2000 ingresaron 3289 pacientes por consulta externa (64,1%), 1839 por urgencias (35,8%) y 5 por nacimientos (0,1%). En el 2005, el 69,4% de los pacientes (4056) ingresaron por consulta externa y el otro 30,6% (1790) por urgencias. Finalmente en el 2010, 4068 pacientes (68,4%) ingresaron por consulta externa, 1876 por urgencias (31,6%) y 2 pacientes ingresaron por el área de nacimientos (0,03%).

#### 4.1.8 Servicio de ingreso

La mayoría de los pacientes con cáncer ingresan a la CCSS por los servicios de Cirugía, Medicina, Gine-obstetricia o Cirugía ambulatoria, tal como se presenta en el siguiente cuadro.

**Cuadro 5: Distribución del número de casos de pacientes con cáncer hospitalizados en la CCSS, según servicio de ingreso.**

Servicio de ingreso	2000		2005		2010	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Cirugía	2,877	56	3,464	59	3,508	59
Medicina	1,076	21	1,090	19	1,079	18
Gine-obstetricia	599	12	725	12	631	11
Cirugía ambulatoria	241	5	341	6	532	9
Otros	340	7	226	4	196	3
<b>Total</b>	<b>5,133</b>	<b>100</b>	<b>5,846</b>	<b>100</b>	<b>5,946</b>	<b>100</b>

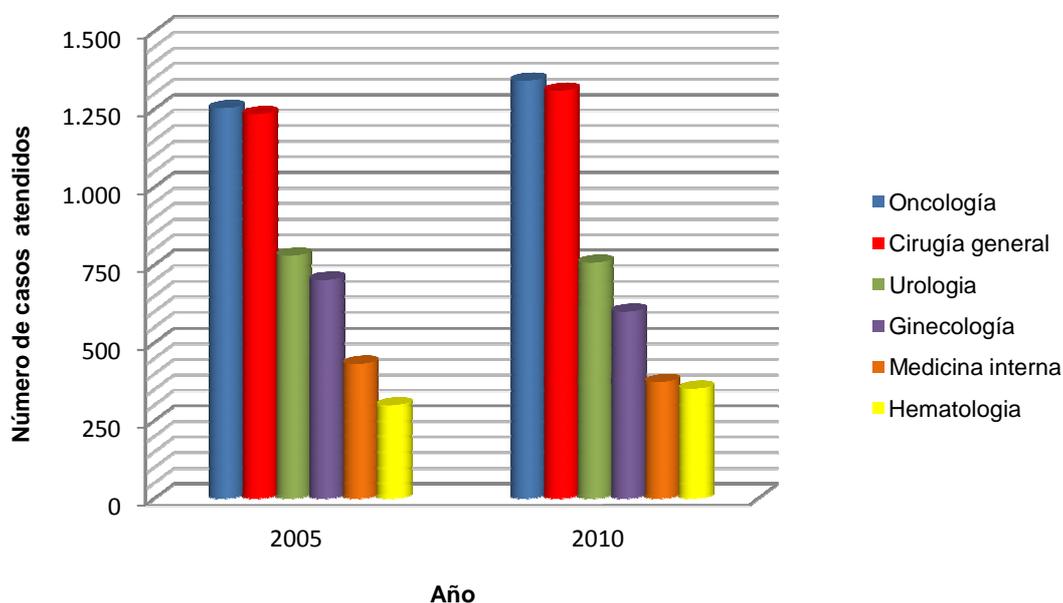
Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

Por el servicio de Cirugía ingresan la mayor cantidad de pacientes; en el año 2000 ingresaron 2877 pacientes (56% del total) mientras que en los años 2005 y 2010 el 59% de los pacientes ingresaron por este servicio (3464 y 3508 casos respectivamente). En el caso del servicio de Medicina, para el año 2000 el 21% de los pacientes ingresaron por este servicio (1076) mientras que para el 2005 y 2010 fueron 19% y 18% respectivamente. Con respecto a Gine-obstetricia y Cirugía ambulatoria en conjunto recibieron a 840 pacientes en el año 2000 (17%), 1066 en el 2005 (18%) y 1163 en el 2010 (20%).

#### 4.1.9 Especialidad de ingreso

En los años 2005 y 2010 el 80% de los pacientes nuevos con cáncer atendidos en la CCSS ingresaron por algunas de las siguientes 6 especialidades de ingreso: Oncología, Cirugía General, Urología, Ginecología, Medicina Interna o Hematología. Del año 2000 no se tienen datos, ya que, en ese período no se registraba la especialidad de ingreso de los pacientes.

**Gráfico 7 : Principales especialidades de ingreso de los pacientes con cáncer hospitalizados en la CCSS en el período 2005-2010.**



Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

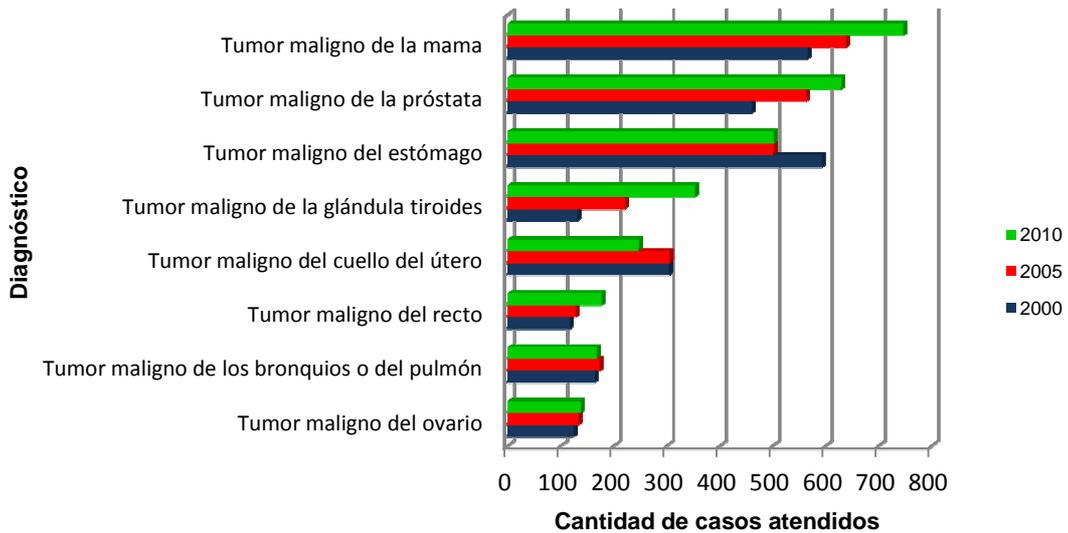
Cómo se observa en el gráfico anterior, los servicios de Oncología y Cirugía general reciben la mayor cantidad de casos; en el 2005 atendieron 1253 y 1234 casos respectivamente, mientras que en el 2010 recibieron 1341 y 1309 pacientes respectivamente. Luego les sigue Urología y Ginecología, con 780 y 702 casos respectivamente en el 2005, y 757 y 600 casos en el 2010. Finalmente, Medicina Interna y Hematología atendieron 432 y 299 pacientes en el año 2005 respectivamente, y 374 y 352 en el 2010.

#### **4.1.10 Diagnóstico principal**

Como se observa en el siguiente gráfico, los 8 tipos de cáncer más comunes diagnosticados en los años 2000, 2005 y 2010 son tumores malignos de: la mama la próstata, el estómago, la glándula tiroides, el cuello del útero, el recto los bronquios o

pulmón y el ovario. Estos representan aproximadamente el 50% del total de casos de pacientes nuevos con cáncer hospitalizados en la CCSS.

**Gráfico 8 : Diagnósticos principales que presentaron la mayor cantidad de pacientes con cáncer hospitalizados en la CCSS en el período 2000-2010.**



Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

En los años 2010 y 2005 el diagnóstico principal fue el cáncer de mamá con 746 y 638 casos respectivamente, luego se ubica el cáncer de próstata con 628 y 563 casos respectivamente, y en tercera posición se encuentra el cáncer de estómago con 502 y 501 casos respectivamente.

En el año 2000 la distribución era diferente, ya que, el diagnóstico principal fue el cáncer de estómago con 593 casos, seguido por el cáncer de mamá con 565 casos y luego el cáncer de próstata con 459 casos.

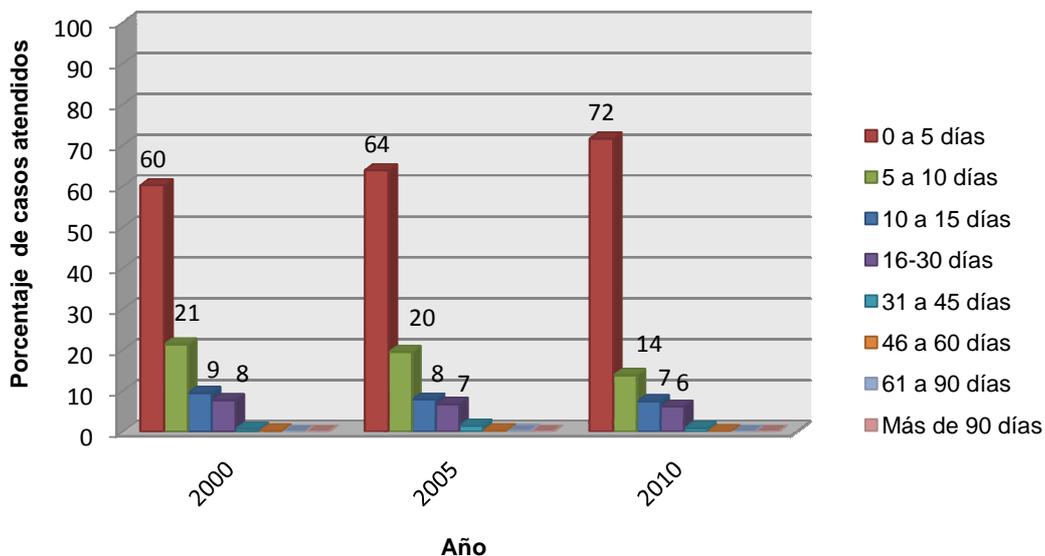
Por otro lado, mientras que en el 2000 y 2005 el cuarto tipo de cáncer más común era el del cuello del útero (con 305 y 306 casos respectivamente) en el 2010 esto cambió y dicha posición la ocupó el cáncer de la glándula tiroides (con 352 casos) y el cáncer de cuello del útero ocupó la quinta posición con 247 casos.

#### 4.1.11 Estancia preoperatoria

La estancia preoperatoria es un indicador clave de la gestión de los centros hospitalarios, ya que, entre más rápido se realice la intervención se evita que la enfermedad avance más rápido, o al menos se le da una mayor calidad de vida y sobrevida al paciente.

En el siguiente gráfico se presenta la distribución del número de casos de pacientes nuevos con cáncer hospitalizados en la CCSS que requirieron cirugía, según la estancia preoperatoria.

**Gráfico 9 : Distribución del número de casos nuevos de pacientes con cáncer hospitalizados que requirieron cirugía, según la estancia preoperatoria para el período 2000-2010.**



Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

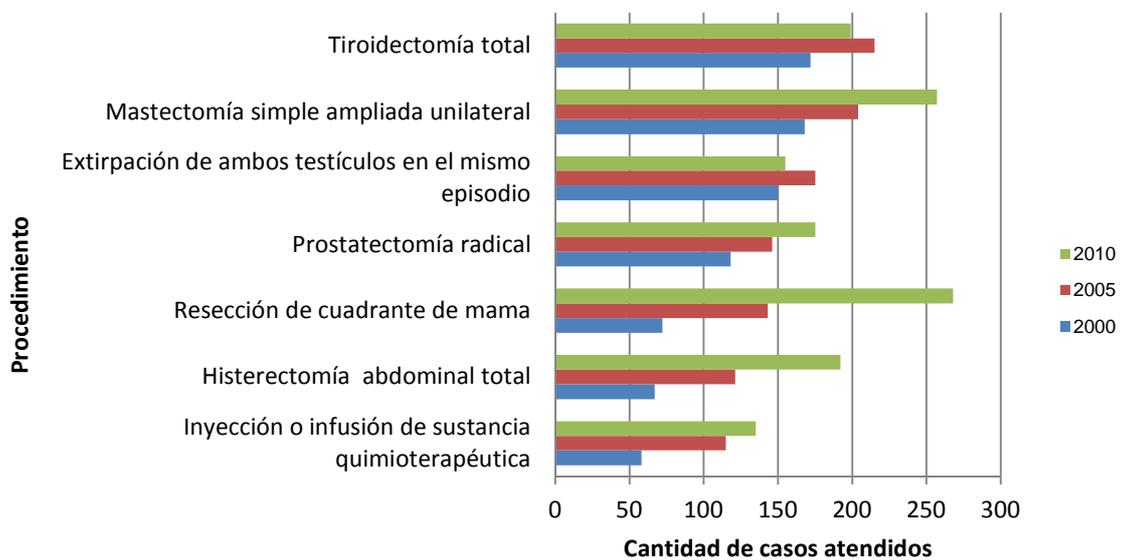
En el año 2000 se le realizó cirugía a 2717 pacientes (53%), en el 2005 a 3373 (58%) y en el 2010 a 5016 pacientes (84%), de los internados por presentar algún tipo de cáncer. Más del 60% de los pacientes que son intervenidos quirúrgicamente esperan menos de 5 días antes de la operación (60% en el 2000, 64% en el 2005 y 72% en el 2010. Por otro lado en los años 2000, 2005 y 2010, el 21%, 20% y 14% del total de pacientes intervenidos respectivamente, presentaron una estancia preoperatoria entre 5 y 10 días.

Adicionalmente entre el 7% y el 9% de los pacientes intervenidos en los años 2000, 2005 y 2010 esperaron entre 10 y 15 días antes de la operación. Menos del 10% del total de pacientes esperaron más de 15 días para ser intervenidos.

#### 4.1.12 Procedimientos desarrollados

En el siguiente gráfico se presentan los procedimientos más comunes que se les realizan a los pacientes con cáncer hospitalizados en la CCSS.

**Gráfico 10 : Principales procedimientos realizados a los pacientes con cáncer hospitalizados en la CCSS en el período 2000-2010.**



Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

Cómo se observa en el gráfico anterior en los años 2000 y 2005 los cuatro tipos de procedimientos que se realizaron con mayor frecuencia fueron: tiroidectomía total (172 y 215 casos respectivamente), mastectomía simple ampliada unilateral (168 y 204 casos respectivamente), extirpación de ambos testículos (150 y 175 casos respectivamente) y prostatectomía radical (118 y 146 casos respectivamente).

En el año 2010 la distribución fue diferente, ya que, el procedimiento que más se realizó fue la resección del cuadrante de mamá (268 casos) luego se ubica la mastectomía simple ampliada unilateral con 257 casos, en tercera posición se encuentra la tiroidectomía total con 199 casos, y después aparece la histerectomía abdominal total con 192 casos.

#### 4.1.13 Condición de salida

Menos del 8,5% de los pacientes que son hospitalizados por primera vez en la CCSS fallecen antes del primer egreso. Como se observa en el siguiente cuadro, aproximadamente el 90% de los pacientes egresan por indicación médica, y entre el 1% y 3% son trasladados a otro centro.

**Cuadro 6: Distribución del número de casos de pacientes con cáncer hospitalizados en la CCSS, según condición de salida.**

Condición de salida	2000		2005		2010	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Indicación médica	4.549	88,6	5.240	89,6	5.339	89,8
Transferido	130	2,5	130	2,2	86	1,4
Exigida	22	0,4	19	0,3	20	0,3
Fuga	3	0,1	3	0,1	2	0,0
Fallecido sin autopsia	366	7,1	378	6,5	444	7,5
Fallecido con autopsia	62	1,2	74	1,3	51	0,9
Fallecido traslado forense	1	0,0	2	0,0	4	0,1
<b>Total</b>	<b>5.133</b>	<b>100</b>	<b>5.846</b>	<b>100</b>	<b>5.946</b>	<b>100</b>

Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

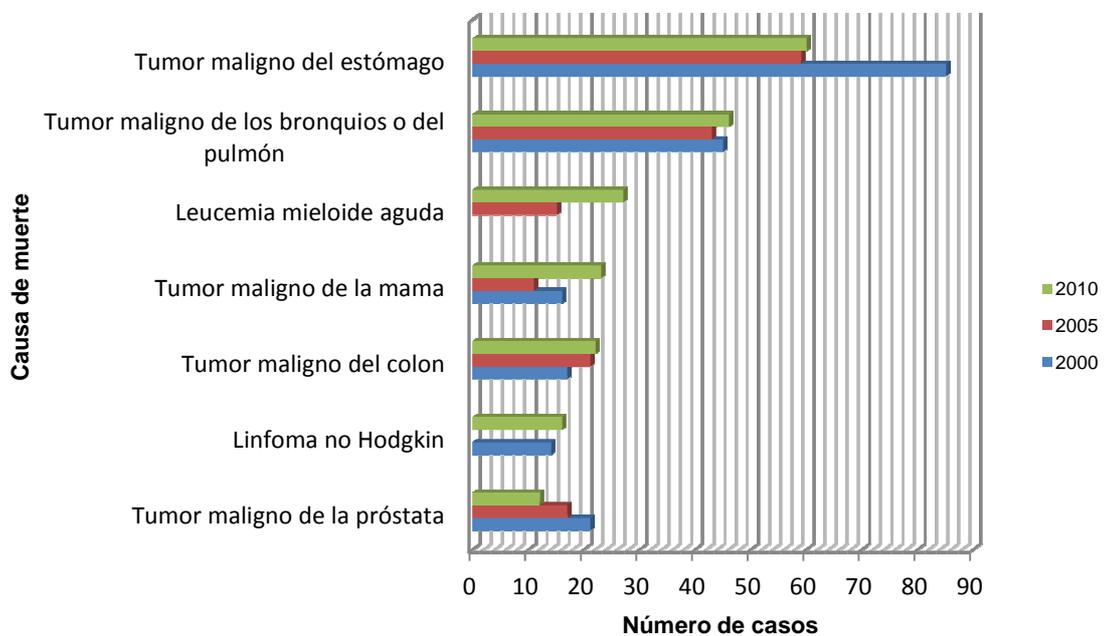
Otras condiciones de salida para los pacientes que no fallecen son la salida exigida, asumiendo el riesgo o por fuga, pero, en ninguno de los tres años de estudio estos pacientes llegan al 1% del total.

Por otro lado, debe indicarse que en el año 2000 fallecieron 429 pacientes, en el 2005 fallecieron 454 y en el 2010 499 pacientes. Del total de egresos, entre el 6,5% y 7,5% egresaron como fallecidos sin autopsia y cerca de 1% como fallecidos con autopsia.

#### 4.1.14 Causa básica de muerte

Dentro del análisis de los casos se puede observar que el de estómago y el de bronquios o pulmón son los dos tipos de cáncer de mayor mortalidad; como se aprecia en el gráfico el tumor maligno de estómago presenta una tendencia al descenso, sin embargo, ocupa el primer lugar desde el año 2000.

**Gráfico 11: Principales causas de muerte de los pacientes con cáncer hospitalizados en la CCSS en el período 2000-2010.**



Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

De los 429 pacientes con cáncer fallecidos en el año 2000, 85 fallecieron por tumor maligno del estómago y 45 por tumor maligno de los bronquios o del pulmón. En el año 2005 fallecieron 454 pacientes, de los cuales 59 tenían cáncer de estómago y 43 cáncer de bronquios o pulmón. Finalmente, de los 500 pacientes fallecidos en 2010, 60 tenían cáncer de estómago y 46 de bronquios o pulmón.

La tercera causa de muerte varió en los tres períodos de estudio. Mientras que en el 2000 fue el tumor maligno de la próstata (con 21 casos), en el 2005 fue el tumor maligno del colon (con 21 casos) el 2010 fue la leucemia mieloide aguda (con 27 casos).

Se observa cómo, aunque el cáncer de mama fue el más común en los períodos analizados, no se ubica entre las tres primeras causas de muerte.

#### 4.1.15 Estancia total

En el siguiente cuadro se presenta la distribución del número de casos de pacientes nuevos con cáncer hospitalizados en la CCSS, según la estancia total de los pacientes en los centros de salud.

**Cuadro 7: Distribución del número de casos de pacientes con cáncer hospitalizados en la CCSS, según estancia total en el período 2000-2010.**

Estancia total	2000		2005		2010	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Menos de 5 días	1,356	28	2,181	37	2,183	37
De 5 a 10 días	1,227	25	1,396	24	1,309	22
De 10 a 15 días	836	17	837	14	828	14
De 15 a 20 días	523	11	466	8	568	10
De 20 a 25 días	317	7	322	6	364	6
De 25 a 30 días	269	6	185	3	242	4
De 30 a 45 días	213	4	293	5	323	5
De 45 a 60 días	77	1.6	99	1.7	88	1.5
De 60 a 90 días	41	0.8	55	0.9	36	0.6
Más de 90 días	8	0.2	12	0.2	5	0.1
<b>Total</b>	<b>5,133</b>	<b>100</b>	<b>5,846</b>	<b>100</b>	<b>5,946</b>	<b>100</b>

Fuente: Base de datos de egresos hospitalarios, Área de Estadísticas en Salud, CCSS.

Como se observa en el cuadro anterior la mayoría de los pacientes presentan una estancia inferior a los 5 días. En el año 2000, el 28% de los pacientes permanecieron menos de 5 días en los centros hospitalarios, mientras que en el 2005 y 2010 este porcentaje ascendió a 37%. Además, 25%, 24% y 22% de los casos presentados en el 2000, 2005 y 2010 respectivamente, permanecieron entre 5 y 10 días hospitalizados.

En el año 2000, 17% de los pacientes permanecieron en los centros de salud entre 10 y 15 días, mientras que en el 2005 y 2010 este porcentaje fue de 14%. Un 24% del total de pacientes presentaron una estancia de entre 15 y 30 días en el año 2000; en esta misma estancia se ubicaron 17% de los pacientes en el 2005 y 20% en el 2010. Los demás casos (aproximadamente 7%) presentaron estancias superiores a los 30 días en los años 2000, 2005 y 2010.

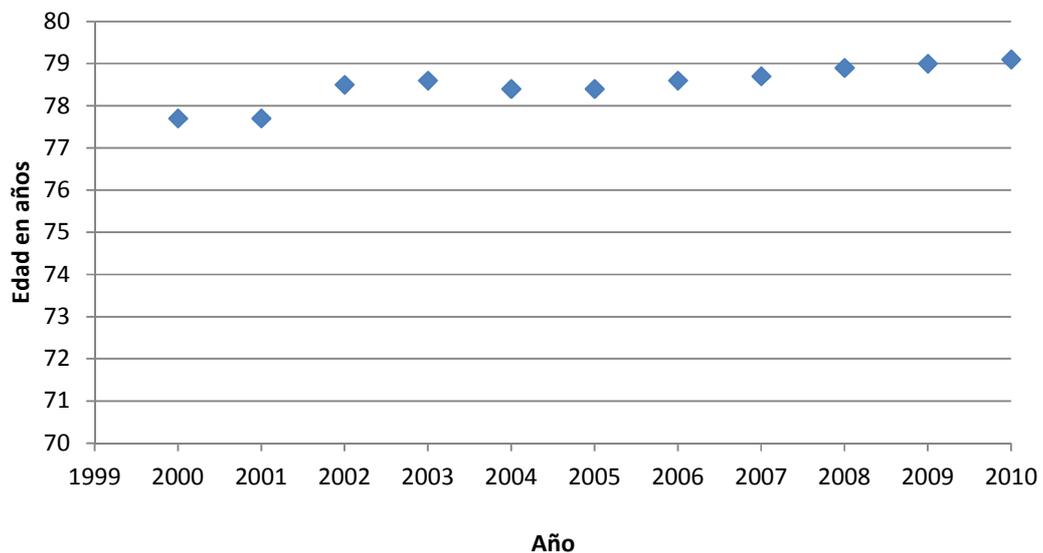
#### 4.1.16 Variables macroeconómicas y sociales que afectan el fenómeno

Dentro de las variables macroeconómicas y sociales que podrían tener una relación con la cantidad de casos nuevos de pacientes hospitalizados con diagnóstico de cáncer en la CCSS, se identificaron las siguientes:

1. *Esperanza de vida al nacer.*

En el gráfico 12 se presenta la evolución de la esperanza de vida al nacer en Costa Rica para el período 2000-2010.

**Gráfico 12: Esperanza de vida al nacer en Costa Rica para el período 2000-2010 (en años).**



Fuente: INEC.

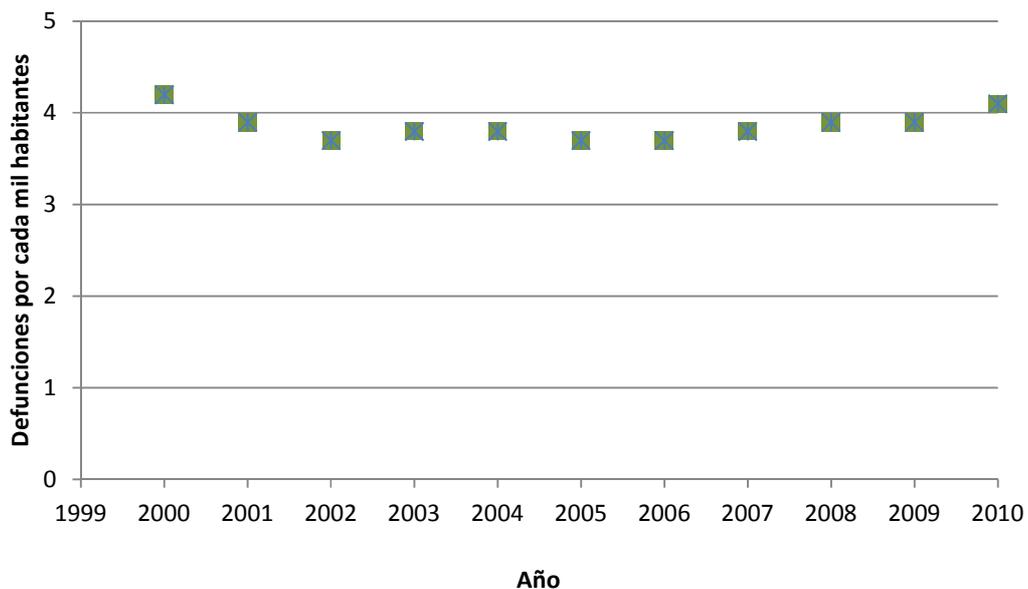
Como se observa en el gráfico anterior, con excepción del período 2003-2004, la tendencia ha sido que la esperanza de vida al nacer vaya creciendo; mientras que en el 2000 se esperaba que una persona viviera aproximadamente 77,5 años, para el 2010 se espera que viva más de 79 años.

## 2. Tasa general de mortalidad.

En los últimos 10 años la tasa de mortalidad en Costa Rica ha sido cercana a 4 defunciones por cada mil habitantes.

En el siguiente gráfico se presenta el comportamiento de la tasa general de mortalidad del país en la última década.

**Gráfico 13: Tasa general de mortalidad en Costa Rica para el período 2000-2010 (en número de defunciones por cada mil habitantes).**



Fuente: INEC.

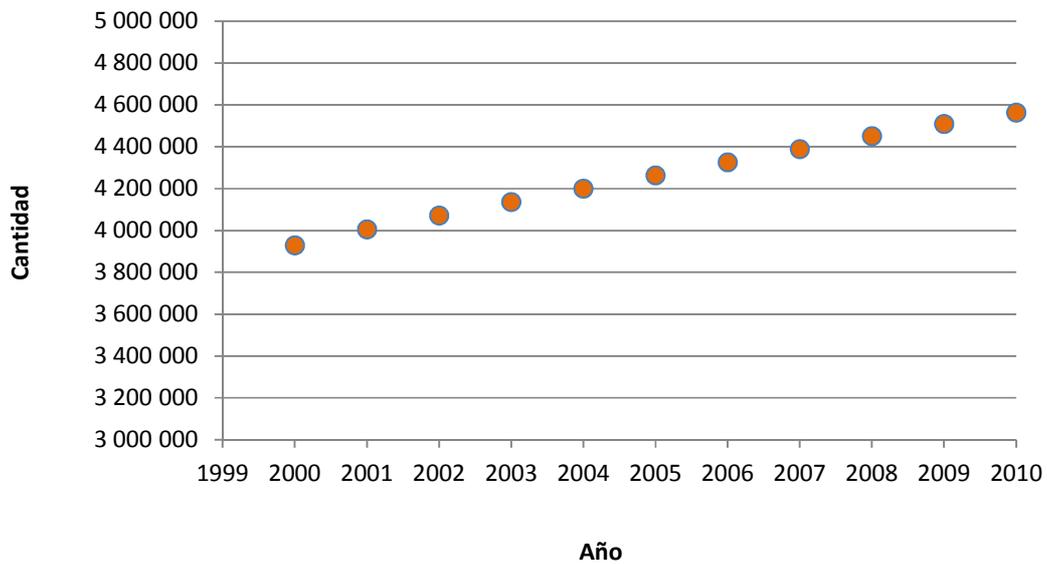
Como se observa en el gráfico 13, en el año 2000 la tasa fue de 4.2 defunciones por cada mil habitantes, a partir de este año se observa una tendencia a disminución hasta llegar a

3.7 en el 2002; posteriormente hay una estabilización en 3.8 para los años 2003 y 2004, y en los últimos 3 años empieza a subir levemente hasta llegar a 4.1 en el 2010.

*Cantidad de Población.*

En el gráfico 14 se presenta la evolución de la cantidad de población que habita en Costa Rica para el período 2000-2010.

**Gráfico 14: Población total de Costa Rica para el período 2000-2010 (en número de personas).**



Fuente: INEC.

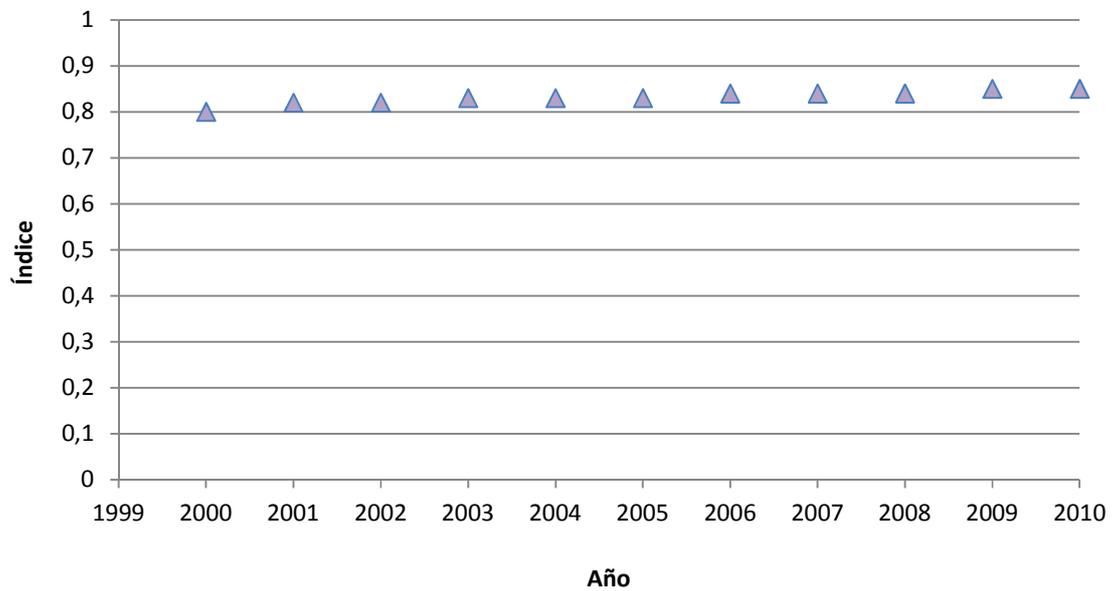
Cómo se observa en el gráfico anterior, la cantidad de población ha estado aumentando prácticamente de forma lineal desde el año 2000, donde la población era de poco más de 3,9 millones, hasta alcanzar una población en el año 2010 de poco menos de 4,6 millones.

### 3. Índice de Desarrollo Humano.

El Índice de Desarrollo Humano (IDH) es un indicador social estadístico compuesto por tres parámetros: vida larga y saludable, educación y nivel de vida digno. Este índice es calculado por el PNUD mediante la aplicación de diferentes instrumentos que miden el nivel de desarrollo de los países de todo el mundo.

En el gráfico 15 se presenta la evolución del IDH de Costa Rica en el período 2000-2010.

**Gráfico 15: Índice de desarrollo humano en Costa Rica para el período 2000-2010.**



Fuente: MIDEPLAN.

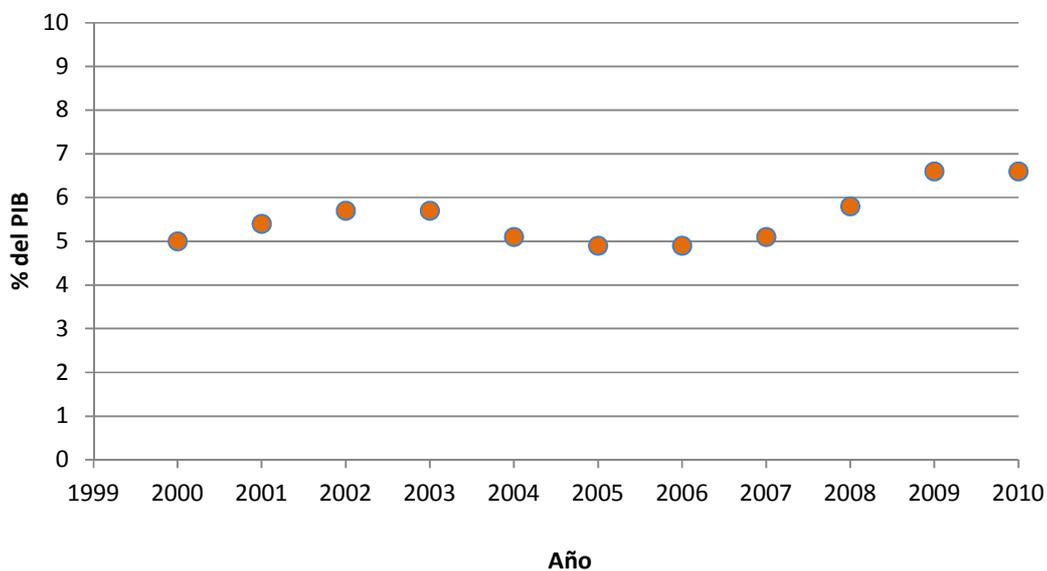
Cómo se observa en el gráfico anterior, el Índice de Desarrollo Humano del país ha tenido un comportamiento muy estable en esta última década, con un muy leve aumento anual que lo lleva desde los 0,8 en el año 2000, hasta 0,84 en el año 2010.

#### 4. Inversión Social en Salud.

Esta variable corresponde al porcentaje del Producto Interno Bruto (PIB) que el país invierte en programas de salud.

En el siguiente gráfico se presenta la inversión social en salud que ha realizado Costa Rica en el período 2000-2010.

**Gráfico 16: Inversión social en salud en Costa Rica para el período 2000-2010 (porcentaje del PIB).**



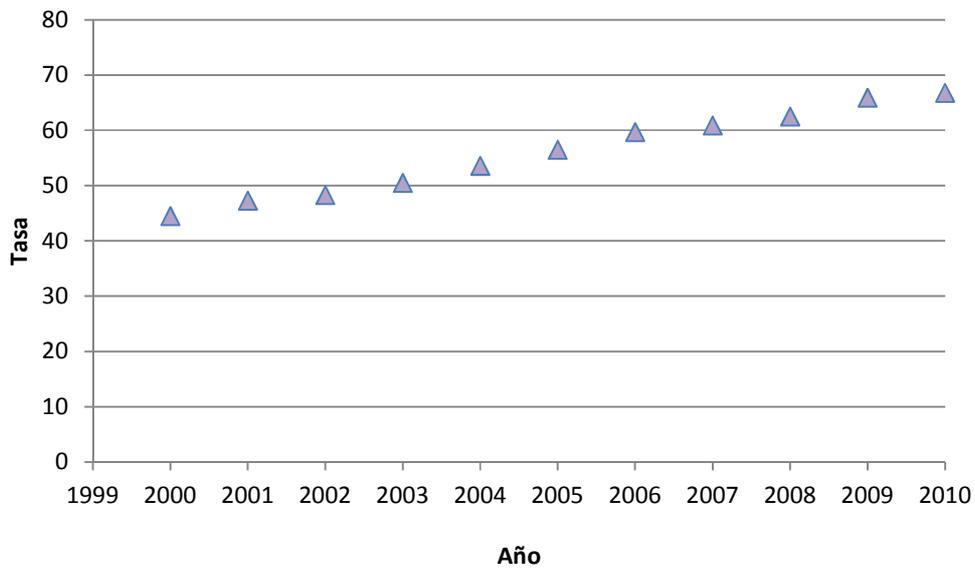
Fuente: MIDEPLAN.

En el gráfico 16 se muestra como la inversión social en salud en el país ha tenido un comportamiento variado, con una leve tendencia al crecimiento, la cual se acentúa del año 2000 al 2003 año en que alcanzó casi 6%; posteriormente se presenta un descenso hasta el 2006 donde llegó a menos de 5%, y de ahí vuelve a subir hasta el 2009 y 2010, donde se alcanzan los valores más altos (casi 7%). Es importante considerar que este aumento puede deberse no tanto al mayor uso de recursos en el tema de salud, sino también a la disminución del PIB producto de la crisis mundial.

5. Tasa de escolaridad y nivel educativo.

La tasa de escolaridad y nivel educativo se refiere al porcentaje de la población del país que tiene concluida la educación diversificada (colegio). En el gráfico 17 se presenta el comportamiento de esta tasa en Costa Rica en el período 2000-2010.

**Gráfico 17: Tasa de escolaridad y nivel educativo en Costa Rica para el período 2000-2010 (porcentaje)**



Fuente: MIDEPLAN.

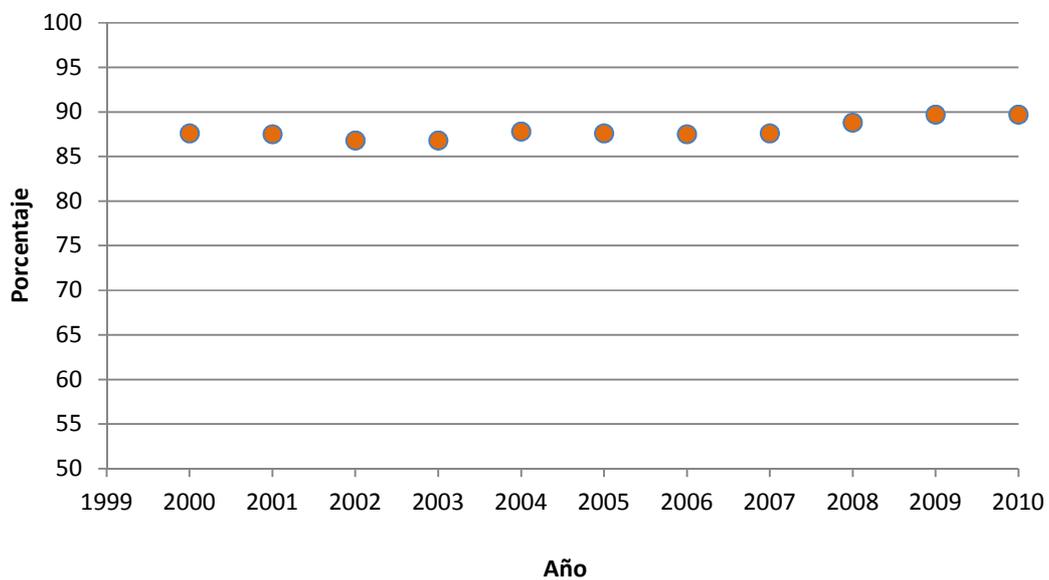
Como se observa en el gráfico anterior, la tasa de escolaridad y nivel educativo ha estado presentando un aumento constante en la última década, pasando de poco más de 40% en el año 2000, a casi 70% en el 2010.

### 6. Cobertura Régimen de Enfermedad y Maternidad.

Corresponde al porcentaje de la población del país que está cubierta por el régimen de enfermedad y maternidad de la CCSS.

En el siguiente gráfico se presenta la evolución de la cobertura del régimen de enfermedad y maternidad en Costa Rica para el período 2000-2010.

**Gráfico 18: Cobertura Régimen de Enfermedad y Maternidad en Costa Rica para el período 2000-2010 (porcentaje de la población nacional).**



Fuente: MIDEPLAN.

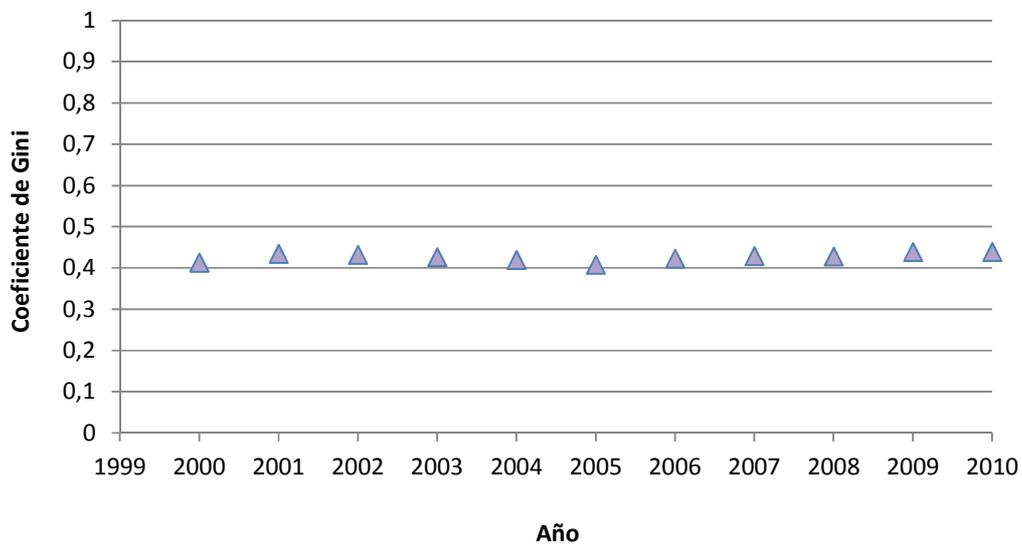
Como se observa en el gráfico 18, en la última década la cobertura del régimen ha presentado un comportamiento prácticamente lineal con valores entre el 85% y 90% de la población nacional.

### 7. Desigualdad en la distribución del ingreso.

Esta variable corresponde a la distribución de los recursos generados en el país entre la población residente. Está medido por el coeficiente de concentración de Gini, el cual tiene un valor entre 0 y 1, donde a medida que el valor se acerca a 1, la concentración de los ingresos se hace mayor

En el siguiente gráfico se presenta la desigualdad en la distribución del ingreso en el país, para los últimos 10 años.

**Gráfico 19: Desigualdad en la distribución del ingreso en Costa Rica para el período 2000-2010 (coeficiente de Gini).**



Fuente: MIDEPLAN.

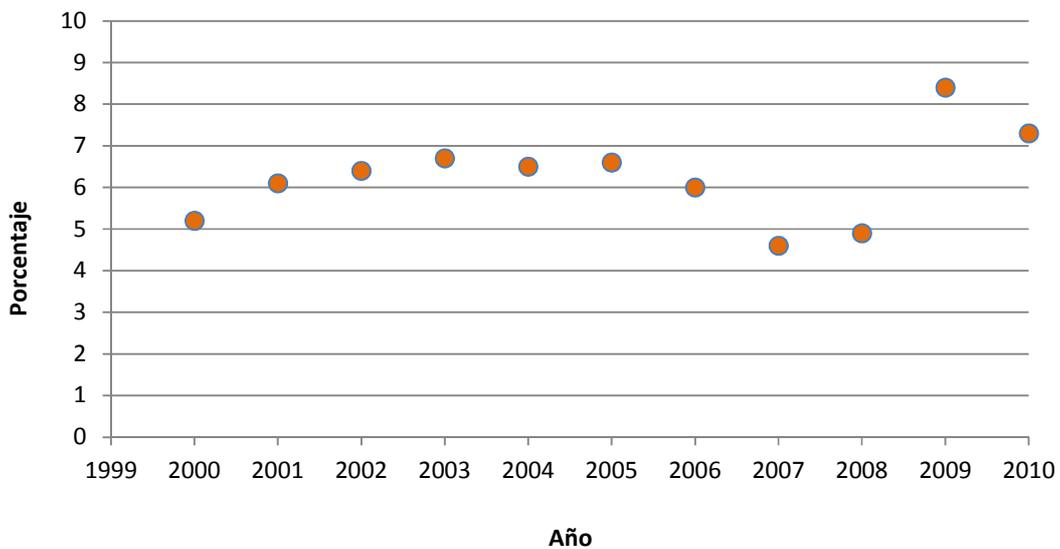
Tal como se observa en el gráfico 19, el comportamiento de esta variable ha sido muy constante en el último decenio, con valores entre los 0,40 y 0,45, para todos los años.

### 8. Tasa de desempleo.

Corresponde al porcentaje de la población económicamente activa que no tiene trabajo.

En el gráfico 20 se presenta la evolución de la tasa de desempleo en Costa Rica durante el período 2000-2010.

**Gráfico 20: Tasa de desempleo en Costa Rica para el período 2000-2010.**



Fuente: MIDEPLAN.

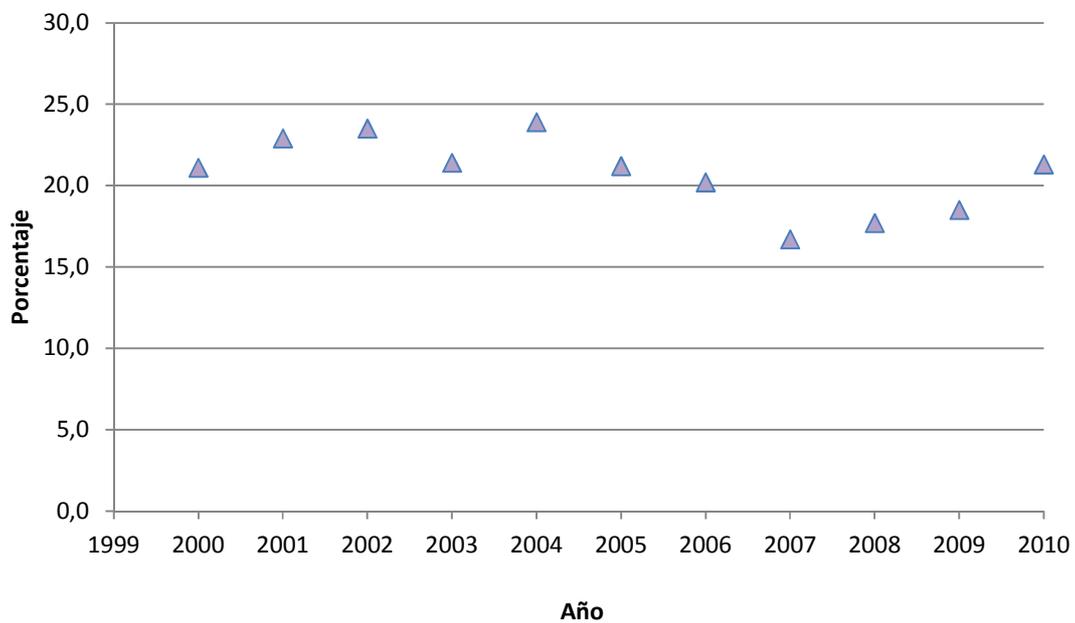
Como se observa en el gráfico anterior, la tasa de desempleo presentó valores relativamente constantes del año 2001 al 2005, cercanos al 7%; posteriormente se observa un descenso en los años 2006 a 2008, llegando a valores menos del 5%; no obstante, para los años 2009 y 2010, producto de la crisis financiera internacional, los valores se elevaron de manera importante hasta estar por encima del 8% en el 2009 y del 7% en el 2010.

### 9. Nivel de pobreza.

Corresponde al porcentaje de la población en la que sus ingresos económicos no alcanzan para suplir las necesidades más básicas de vida, es decir, que se encuentran por debajo del límite de pobreza definido por el Banco Mundial, que es aproximadamente US\$2 diarios.

En el siguiente gráfico se presenta el nivel de pobreza del país en la última década

**Gráfico 21: Nivel de pobreza en Costa Rica para el período 2000-2010.**



Fuente: MIDEPLAN.

Tal como se observa en el gráfico 21, entre los años 2000 y 2006, el nivel de pobreza mostró porcentajes contantes con valores entre el 20% y el 24%, posteriormente en el año 2007 presentó descenso importante hasta un valor inferior al 17%, pero, posteriormente del 2008 al 2010 volvió a aumentar hasta un valor superior al 21% para el año 2010.

## **4.2 Selección de las variables de entrada**

De acuerdo a la finalidad de la red neuronal, se deben identificar las variables de entrada, es decir, aquellas que podrían afectar el comportamiento de las variables dependientes, que constituyen la salida de la red.

En esta investigación se desarrollaron dos modelos de redes neuronales diferentes, en el que cada uno tiene una salida propia:

- La red de casos nuevos de pacientes con diagnóstico de cáncer hospitalizados en la CCSS (en adelante “Red de casos nuevos”), cuya salida es el número de casos esperados en el mes.
- La red de estancia hospitalaria de los pacientes con diagnóstico de cáncer (en adelante “Red de estancia”), cuya salida es el número esperado de días de hospitalización del paciente a partir de su ingreso.

### **4.2.1 Variables de entrada para la red de casos nuevos**

La salida de esta red es mensual y no depende de las características individuales de los pacientes con cáncer ni de las condiciones de los centros de salud de la C.C.S.S (datos a los que se tiene acceso en la base de datos de egresos hospitalarios de la C.C.S.S).

Una de las bondades de las redes neuronales, es que permite identificar relaciones entre variables de entrada que permiten modelar de mejor manera una variable de salida. Es decir, si no se dispusiera de variables de entrada probablemente sería más recomendable utilizar una técnica diferente a las redes neuronales.

Tal como se ha descrito anteriormente, debido a la naturaleza de la salida de esta red, fue necesario buscar información sobre variables macro del país, de orden económico y social, que podrían incidir en la cantidad de casos nuevos de cáncer que surgen mensualmente.

Las variables identificadas y cuantificadas para este fin, fueron descritas en el apartado 1.1.16, las cuales son:

- *Esperanza de vida al nacer.*
- *Tasa general de mortalidad.*
- *Cantidad de Población.*
- *Índice de Desarrollo Humano.*
- *Inversión Social en Salud.*
- *Tasa de escolaridad y nivel educativo.*
- *Cobertura Régimen de Enfermedad y Maternidad.*
- *Desigualdad en la distribución del ingreso.*
- *Tasa de desempleo.*
- *Nivel de pobreza.*

Debe mencionarse que las variables anteriores podrían no afectar linealmente al fenómeno de estudio, es decir los valores de estas variables para un año determinado podrían no tener incidencia en el número de casos nuevos en ese mismo año, sino uno, dos o más años después. Por lo que, como parte de la selección de las variables para esta red, debió determinarse cuál es el período que tiene mayor efecto sobre el fenómeno de estudio.

Por su parte, de acuerdo con una entrevista realizada a un profesor de la Universidad de Costa Rica, experto en el tema, se determinó la necesidad de incluir en la red neuronal variables relacionadas con el comportamiento en el tiempo del fenómeno de estudio. Por lo que además de las mencionadas anteriormente, se consideraron las siguientes:

1. Año.
2. Mes.
3. Promedio de los últimos 3 años.

4. Promedio de los últimos 2 años.
5. Promedio del último año.
6. Promedio de los últimos 6 meses.
7. Promedio de los últimos 3 meses.

Estas últimas 5 variables se incorporaron como variables de entrada, ya que, contribuyen a que el aprendizaje de la red neuronal sea más rápido y efectivo.

En total, como base de partida para seleccionar las variables de entrada a incorporar en la red, se dispone de 17 variables.

Para definir cuales variables incluir en la red se utilizaron dos programas informáticos estadísticos: SPSS y Statistica, ambos tienen un módulo sobre redes neuronales, en el cual, de acuerdo al comportamiento de los datos, recomiendan que tipo de red se debe utilizar, cuantas neuronas ocultas utilizar e indican el error esperado para las redes.

Primeramente, se modeló en Statistica la red, utilizando las 17 variables de entrada disponibles, pero, variando el período de incidencia de las variables macro. Es decir se evaluó la red utilizando un rezago de las variables macro a un año, a tres años y a 5 años.

El resultado de dicho análisis se observa en el Cuadro 8, donde se determina que las redes neuronales con mejores resultados (mayor efectividad y menor error), correspondían a aquellas generadas utilizando un rezago en las variables macro de tres años.

**Cuadro 8: Determinación de período de incidencia de variables macro, con Statistica.**

Escenario de rezago	Efectividad promedio	Error promedio
Un año	83.14%	0.51%
Tres años	87.42%	0.39%
Cinco años	86.80%	0.41%

Fuente: Los Autores

Posteriormente se realizó un análisis de correlaciones con el fin de determinar si existe dependencia entre las mismas, de modo que se justifique su incorporación en el modelo de redes neuronales.

Para realizar el análisis se utilizó el índice de correlación de Pearson, ya que, se trata de relaciones entre variables cuantitativas. Este se desarrolló mediante el programa estadístico SPSS, generando los siguientes resultados:

**Cuadro 9: Análisis de correlaciones entre las variables de entrada para la Red de casos nuevos, con SPSS.**

Correlaciones utilizando Coeficiente de Pearson																		
Variables	Casos	Año	Mes	Promedio 3 años	Promedio 2 año	Promedio 1 año	Promedio 6 meses	Promedio 3 meses	Esperanza de vida al nacer	Tasa general de mortalidad	Cantidad de Población	Indice de Desarrollo Humano	Inversión Social en Salud (% del PIB)	Tasa de escolaridad y nivel educativo*	Cobertura Regimen de Enfermedad y Maternidad	Desigualdad en la distribución del ingreso**	Tasa de desempleo	Nivel de pobreza
Casos																		
Año																		
Mes																		
Promedio 3 años																		
Promedio 2 año																		
Promedio 1 año																		
Promedio 6 meses																		
Promedio 3 meses																		
Esperanza de vida al																		
Tasa general de																		
Cantidad de Población																		
Indice de Desarrollo																		
Inversión Social en																		
Tasa de escolaridad y																		
Cobertura Regimen de																		
Desigualdad en la																		
Tasa de desempleo																		
Nivel de pobreza																		

Existe correlación  
 No existe correlación

Fuente: SPSS

Como se observa en el cuadro anterior, las siguientes 7 variables no tienen una correlación significativa con la variable de salida del modelo (número de casos): año, inversión social en salud, tasa de escolaridad y nivel educativo, esperanza de vida al nacer, tasa general de mortalidad, cantidad de población, y nivel de pobreza. Estas últimas 4, tampoco tienen dependencia con otras variables relaciones directamente con la salida, como lo son el promedio de casos de los últimos años y meses.

Por otro lado, las variables año y mes tienen relación con muy pocas de las restantes 15 variables, sin embargo, para efectos del aprendizaje del modelo, debe incluirse al menos una variable temporal. Se seleccionó la variable mes.

Finalmente para definir cuales variables se debían descartar y cuales se debían incluir, se modeló la red en SPSS con las 17 variables. Como parte de los resultados, este programa informático muestra la importancia relativa de las diferentes variables en la red neuronal. Lo cual permitió hacer la siguiente priorización:

**Cuadro 10: Priorización de variables de entrada Red de casos nuevos, con SPSS.**

Variables	Importancia	Importancia normalizada
Promedio 6 meses	0,14	100,0%
Promedio 2 año	0,11	80,8%
Tasa de desempleo	0,08	60,3%
Promedio 3 años	0,08	58,0%
Cobertura Régimen de EM	0,07	52,9%
Nivel de pobreza	0,06	46,0%
Promedio 1 año	0,06	42,3%
Mes	0,05	40,0%
Promedio 3 meses	0,05	39,7%
Índice de Desarrollo Humano	0,05	38,9%
Tasa general de mortalidad	0,05	36,5%
Inversión Social en Salud (% del PIB)	0,04	26,6%
Tasa de escolaridad y nivel educativo	0,04	26,4%
Desigualdad en la distribución del ingreso	0,03	22,3%
Esperanza de vida al nacer	0,03	21,9%
Año	0,03	20,3%
Cantidad de Población	0,02	17,3%

Fuente: Los Autores

Como se observa en el cuadro anterior las variables: inversión social en salud, tasa de escolaridad y nivel educativo, desigualdad en la distribución del ingreso, esperanza de vida al nacer, año y cantidad de población presentan una importancia normalizada menor al 30%, por lo que se determinó que no se debían incorporar al modelo de red neuronal, con excepción de la inversión social en salud, que si se incorporó debido a su relación directa con el objeto de estudio.

De esta forma, las variables de entrada para construir la red neuronal de casos nuevos de pacientes con cáncer son las siguientes:

- |                               |                                |
|-------------------------------|--------------------------------|
| 1. Mes                        | 8. Índice de Desarrollo Humano |
| 2. Promedio 3 años            | 9. Inversión Social en Salud   |
| 3. Promedio 2 años            | 10. Cobertura Régimen de       |
| 4. Promedio 1 año             | Enfermedad y Maternidad        |
| 5. Promedio 6 meses           | 11. Tasa de desempleo          |
| 6. Promedio 3 meses           | 12. Nivel de pobreza           |
| 7. Tasa general de mortalidad |                                |

#### **4.2.2 Variables de entrada para la red de estancia**

Para la red de estancia, las variables que afectan la cantidad de días que un paciente con cáncer permanece hospitalizado corresponden a características y condiciones del paciente y del servicio de salud en que es atendido.

En ese sentido, para determinar las variables de entrada para la red de estancia, se tomó como base de partida la base de datos de egresos hospitalarios de la C.C.S.S, la cual es administrada por el Área de Estadísticas en Salud de dicha institución.

Esta base está constituida por un total de 98 variables, pero, no todas inciden en la cantidad de días de hospitalización de los pacientes con cáncer. Considerando que la finalidad de la red de estancia es poder pronosticar la cantidad de días que va estar hospitalizado un paciente con cáncer, a partir de la fecha de ingreso y conociendo su diagnóstico principal, como posibles variables de entrada a la red se tienen las siguientes.

- |                            |                          |
|----------------------------|--------------------------|
| 1. Centro de Salud         | 9. Diagnóstico principal |
| 2. Sexo                    | 10. Diagnóstico 2        |
| 3. País de nacimiento      | 11. Diagnóstico 3        |
| 4. Residencia              | 12. Diagnóstico 4        |
| 5. Área de ingreso         | 13. Diagnóstico 5        |
| 6. Servicio de ingreso     | 14. Diagnóstico 6        |
| 7. Especialidad de ingreso | 15. Edad del paciente    |
| 8. Ingresos previos        |                          |

De estas 15 variables, se descartaron los diagnósticos adicionales (del 2 al 6) del paciente, conservando únicamente el principal, ya que, no en todos los casos se disponía de información sobre seis diagnósticos, por lo que si se incluían afectaría la consistencia de la red. De igual forma no se disponía de información de todos los casos sobre la especialidad de ingreso, por lo que esta variable fue descartada. Además, el país de nacimiento se excluyó, ya que, más del 90% de los casos corresponden a pacientes nacidos en Costa Rica, por lo que esta variable no era significativa para el modelo.

De esta forma quedaron 8 variables posibles a incorporar en la red de estancia:

1. Edad del paciente
2. Sexo
3. Residencia
4. Centro de Salud

5. Área de ingreso
6. Servicio de ingreso
7. Ingresos previos
8. Diagnóstico principal

Posteriormente, se realizó un análisis de correlaciones entre las variables anteriores con el fin de determinar si existe dependencia entre las mismas, de modo que se justifique su incorporación en el modelo de redes neuronales.

Para realizar el análisis se utilizó el estadístico *Chi* cuadrado para las relaciones entre variables categóricas, los estadísticos *F* y *D* de Cohen para las relaciones entre una variable categórica y una variable cuantitativa, y el índice de Pearson para relaciones entre variables cuantitativas. Para calcular los diferentes índices se utilizó el programa estadístico SPSS, generando los siguientes resultados:

**Cuadro 11: Análisis de correlaciones entre las variables de entrada para la Red de estancia, con SPSS.**

Variables	Área de ingreso			Centro de salud			Diagnostico principal			Residencia			Servicio de ingreso			Sexo			Edad del egreso			Ingresos anteriores			Estancia del paciente					
	Estadístico	Valor	Significancia	Estadístico	Valor	Significancia	Estadístico	Valor	Significancia	Estadístico	Valor	Significancia	Estadístico	Valor	Significancia	Estadístico	Valor	Significancia	Estadístico	Valor	Significancia	Estadístico	Valor	Significancia	Estadístico	Valor	Significancia			
Área de ingreso				CHI	3194,2	0	CHI	8465,0	0	CHI	867,5	0	CHI	10377,5	0	CHI	637,0	0	F Cohen	0,0335	0	F Cohen	0,0068	0	F Cohen	0,1519	0	F Cohen	0,1971	0
Centro de salud							CHI	30794,8	0	CHI	336092,8	0	CHI	76376,2	0	CHI	622,6	0	F Cohen	0,4629	0	F Cohen	0,2715	0	F Cohen	0,1971	0	F Cohen	0,329	0
Diagnostico principal										CHI	8315,7	0	CHI	106053,7	0	CHI	37041,6	0	F Cohen	0,6264	0	F Cohen	0,2228	0	F Cohen	0,329	0	F Cohen	0,1613	0
Residencia													CHI	5771,4	0	CHI	394,8	0	F Cohen	0,1184	0	F Cohen	0,2175	0	F Cohen	0,3048	0	F Cohen	0,2362	0
Servicio de ingreso																CHI	9599,9	0	F Cohen	0,4979	0	F Cohen	0,1448	0	F Cohen	0,3048	0	F Cohen	0,029	0
Sexo																			D Cohen	0,3125	0	D Cohen	0,086	0	D Cohen	0,2362	0	D Cohen	0,029	0
Edad de egreso																						Pearson	-0,148	0	Pearson	0,029	0	Pearson	-0,069	0
Ingresos anteriores																												Pearson	-0,069	0
Estancia del paciente																														

Existe dependencia  
No existe dependencia

Fuente: SPSS

Cómo se observa en el cuadro anterior, con un nivel de confianza del 90% hay una dependencia significativa<sup>1</sup> entre todas las variables independientes (de entrada) con la variable dependiente (salida = estancia). Asimismo hay dependencia entre todas las

<sup>1</sup> En el caso del programa SPSS, la significancia igual a cero indica que hay dependencia; conforme la significación se acerca o se aleja de cero será la intensidad de esa dependencia.

variables de entrada disponibles, aunque entre algunas esta dependencia es menos intensa.

Con base en el análisis anterior, se decidió trabajar con las 8 variables de entrada disponibles para modelar la red neuronal de estancia (centro de salud, sexo, residencia, área de ingreso, servicio de ingreso, ingresos previos, diagnóstico principal y edad del paciente) ya que, cada una puede aportar elementos diferentes que favorezcan el aprendizaje de la red.

### **4.3 Diseño del modelo de redes neuronales**

*“Las RNA son modelos matemáticos que tratan de hacer una pequeña emulación del funcionamiento del cerebro humano”* (Toro, Mejía y Salazar, 2004, p.26). Estas redes se aprenden la dinámica del fenómeno en estudio y tienen en cuenta el tiempo como una variable que afecta el fenómeno. Para su diseño se deben establecer varias características.

#### **4.3.1 Tipo de red**

Según el tipo de arquitectura y funciones que se utilicen, se tienen diversos modelos de redes neuronales artificiales. Para el desarrollo de aplicaciones de predicción de datos, el modelo más utilizado es el de perceptrón multicapa con un aprendizaje de propagación hacia atrás. (Cevalloa, 2004).

En el caso de la presente investigación, tanto para la red de casos nuevos como para la red de estancia, se utilizó el modelo perceptrón multicapa con propagación hacia atrás, ya que, la finalidad es desarrollar aplicaciones para realizar proyecciones en el tiempo.

#### **4.3.2 Estructuración de la red**

Se le llama arquitectura de una red neuronal a la estructura de interconexión del conjunto de nodos que componen la red, y que se organizan en capas.

Para el desarrollo de las redes de neuronales de casos nuevos y de estancia se consideraron tres capas: una de entrada, una intermedia y una de salida.

El número de nodos en la capa de entrada depende de las variables consideradas para establecer un patrón entre la representación de la entrada y la de la salida.

En el caso de la red de casos nuevos, como se indicó en la sección 1.2 de este capítulo, se seleccionaron 12 variables de entrada, las mismas son cuantitativas y por consiguiente para la estructuración de la red se consideraron 12 nodos o neuronas de entrada.

Con respecto a la red de estancia, se seleccionaron 8 variables de entrada, de las cuales 6 son variables categóricas y 2 cuantitativas.

De acuerdo con lo que definen algunos teóricos, cuando las variables de entrada a una red neuronal son categóricas, el número de neuronas de la capa entrada depende del número de categorías, es decir se debería considerar una neurona para cada categoría. No obstante, para efectos de la red de estancia que se desarrolló en esta investigación, si se seguía dicha recomendación el número de neuronas de entrada superaría las 500, ya que, por ejemplo, una de las variables disponía de 80 categorías; esto le habría restado sencillez al modelo y habría generado limitaciones a nivel del programa informático utilizado.

Considerando lo anterior, lo que procedió fue darles carácter cuantitativo a las variables categóricas asignando un valor numérico a cada categoría, de esta manera el número de neuronas establecido para la capa de entrada en la red de estancia fue de 8.

Para asegurarse de que la consideración anterior no afectaba los resultados generados por la red neuronal, se realizaron varias pruebas con diferentes valores numéricos para cada categoría, logrando determinar que los errores de pronóstico no presentaban variaciones significativas al realizar la cuantificación indicada.

La cantidad de nodos en la capa de salida depende del número de salidas que tenga el modelo; si la variable de salida es categórica debe contemplarse una neurona o nodo por cada categoría.

Cómo se indicó anteriormente, cada una de las redes neuronales objeto de la presente investigación tienen una única de salida y por consiguiente esta capa tiene solo una neurona o nodo, ya que, ambas son cuantitativas: casos nuevos de pacientes con cáncer por mes (para la red de casos nuevos) y días de estancia u hospitalización del paciente (para la red de estancia).

El número de nodos en la capa intermedia varia estableciéndose en muchos casos mediante prueba y error para elegir el número más conveniente. En el caso de las redes de casos nuevos y de estancia, para determinar la cantidad de neuronas en la capa intermedia o capa oculta se utilizó los paquetes informáticos SPSS y Statistica.

Ambas redes se modelaron en los 2 programas mencionados, generando los siguientes resultados.

**Figura 4: Resultados del modelado de la red de casos nuevos con Statistica**

Summary of active networks (RNA Incidencia 3 año diferencia sta)						
Index	Net. name	Training perf.	Test perf.	Training error	Test error	Training algorithm
1	MLP 12-6-1	0,659123	0,816941	0,007398	0,009977	BFGS 21
2	MLP 12-7-1	0,733835	0,841305	0,006062	0,007498	BFGS 44
3	MLP 12-7-1	0,765509	0,873585	0,005430	0,006416	BFGS 39
4	MLP 12-8-1	0,757084	0,821249	0,005571	0,008577	BFGS 40
5	MLP 12-5-1	0,816502	0,842384	0,004351	0,006669	BFGS 60

Fuente: Statistica

**Figura 5: Resultados del modelado de la red de casos nuevos con SPSS**

Información sobre la red			
Capa de entrada	Covariables	1	Promedio 3 años
		2	Promedio 2 año
		3	Promedio 1 año
		4	Promedio 6 meses
		5	Promedio 3 meses
		6	Tasa general de mortalidad
		7	Índice de Desarrollo Humano
		8	Inversión Social en Salud (% del PIB)
		9	Cobertura Regimen de Enfermedad y Maternidad
		10	Tasa de desempleo
		11	Nivel de pobreza
		12	Mes
	Número de unidades <sup>a</sup>		12
	Método de cambio de escala de las covariables		Tipificados
Capas ocultas	Número de capas ocultas		1
	Número de unidades de la capa oculta 1 <sup>a</sup>		7
Capa de salida	Función de activación		Tangente hiperbólica
	Variables dependientes	1	Casos
	Número de unidades		1
	Método de cambio de escala para dependientes de escala		Tipificados
	Función de activación		Identidad
	Función de error		Suma de cuadrados

a. Sin incluir la unidad de sesgo

Fuente: SPSS

**Figura 6: Resultados del modelado de la red de estancia con Statistica**

Summary of active networks (Spreadsheet2 in Datos entrada finales red estancia Stat)						
Index	Net. name	Training perf.	Test perf.	Training error	Test error	Training algorithm
1	MLP 8-6-1	0.455565	0.402644	0.009521	0.010242	BFGS 10
2	MLP 8-11-1	0.439607	0.394583	0.009719	0.010329	BFGS 8
3	MLP 8-8-1	0.459027	0.403285	0.009481	0.010245	BFGS 11
4	MLP 8-7-1	0.455267	0.400326	0.009524	0.010263	BFGS 9
5	MLP 8-4-1	0.459272	0.401914	0.009478	0.010257	BFGS 10

Fuente: Statistica

**Figura 7: Resultados del modelado de la red de estancia con SPSS**

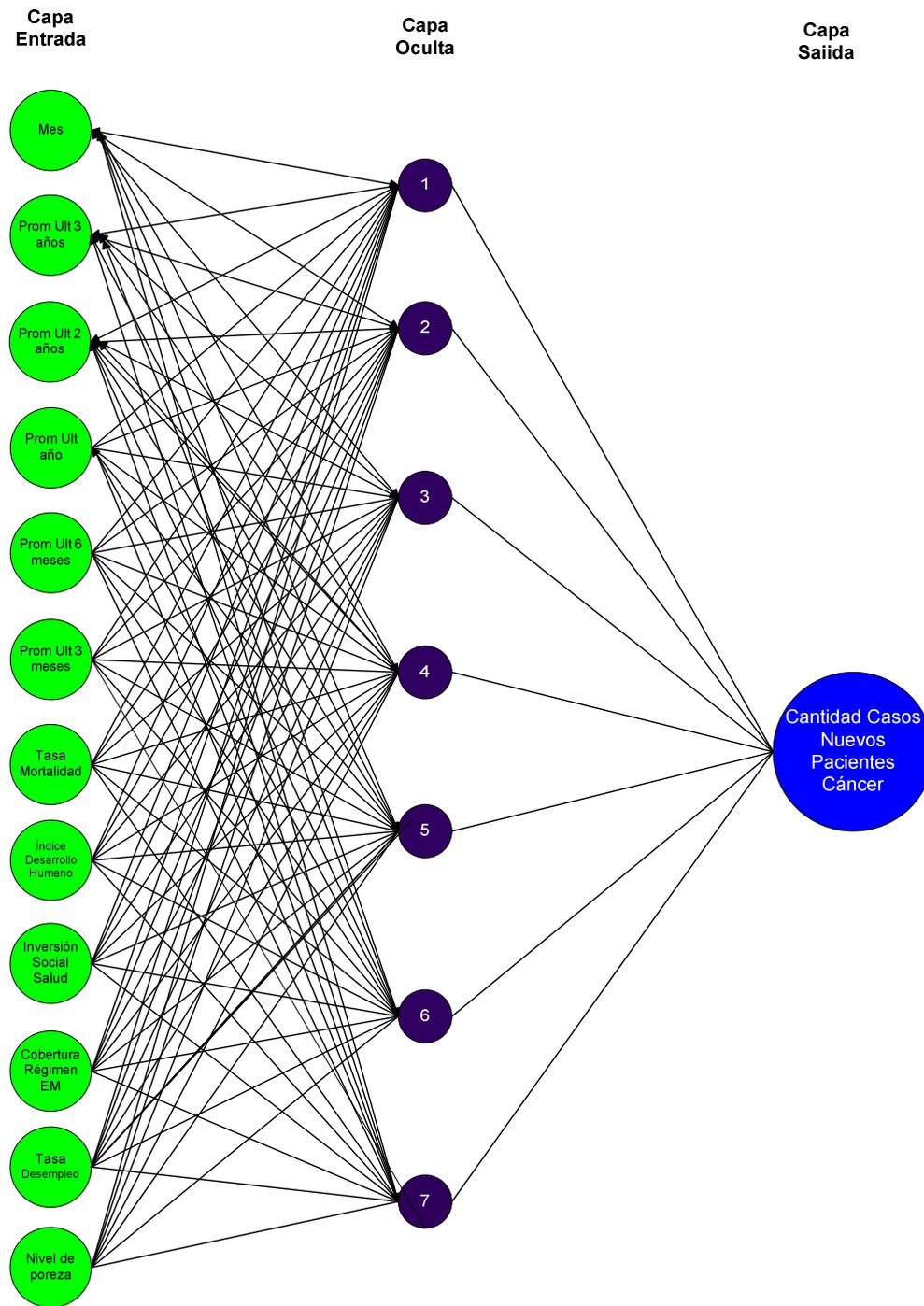
Network Information			
Input Layer	Covariates	1	Edad
		2	Sexo
		3	Residencia
		4	Centro de Salud
		5	Área de ingreso
		6	Servicio de ingreso
		7	Ingresos anteriores
		8	Diagnostico principal
	Number of Units <sup>a</sup>		8
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1 <sup>a</sup>		7
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Estancia del paciente
	Number of Units		1
	Rescaling Method for Scale Dependents		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares

Fuente: SPSS

Cómo se observa en las figuras anteriores, en el caso de la red de estancia, el modelo de la red en SPSS recomienda utilizar 7 neuronas ocultas. Por lo que considerando que una de las cinco mejores redes generadas con Statistica también utiliza 7 neuronas ocultas, se definió utilizar esta cantidad para la red de estancia.

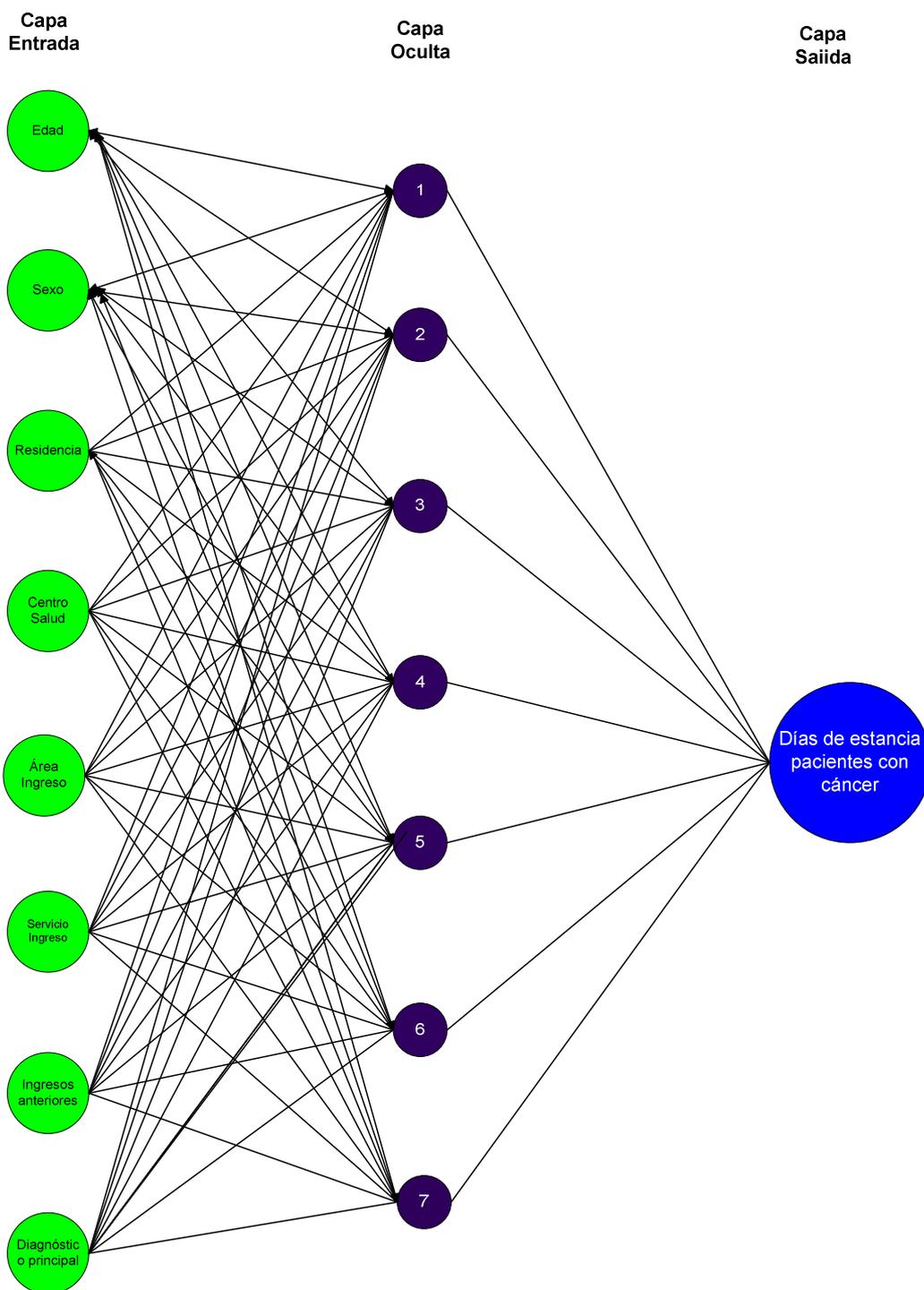
En el caso de la red de casos nuevos se definió la utilización de 7 neuronas ocultas, ya que, al analizar los resultados obtenidos con Statistica se observa que este número de variables ocultas es de los que generan mejores resultados; adicionalmente, del análisis realizado con SPSS se recomienda utilizar este número de variables ocultas. En las siguientes figuras se presenta la estructuración de las redes neuronales de casos nuevos y de estancia para esta investigación:

Figura 8: Estructuración de la red neuronal de casos nuevos de pacientes con cáncer.



Fuente: Los autores

Figura 9: Estructuración de la red neuronal de estancia de los pacientes con cáncer.



Fuente: Los autores

### 4.3.3 Funcionamiento de la red

Cada neurona de la capa de entrada está caracterizada en cualquier instante por un valor numérico denominado valor o estado de activación  $a_i(t)$ . Este corresponde al valor de cada variable de entrada que alimenta la red.

**Cuadro 12: Ejemplo de neuronas de entrada en la Red de casos nuevos**

Neurona de entrada	Valor o estado de activación $a_i(t)$
Mes	5 (mayo)
Promedio 3 años	527 casos nuevos
Promedio 2 año	527 casos nuevos
Promedio 1 año	527 casos nuevos
Promedio 6 meses	527 casos nuevos
Promedio 3 meses	509 casos nuevos
Tasa general de mortalidad	4
Índice de Desarrollo Humano	0,85
Inversión Social en Salud	4,8%
Cobertura Régimen de Enfermedad y Maternidad	86,2%
Tasa de desempleo	4,2%
Nivel de pobreza	20%

Fuente: Los autores

A estos valores de entrada se les aplica una función de activación para generar señales hacia las neuronas ocultas. En el caso de las dos redes objeto de esta investigación se definió que la función de activación sería la función identidad  $F(x) = x$ , por lo que el parámetro que se le pasa a la neurona oculta será directamente el valor de la variable de entrada ( $a_i(t) = x_i$ ) multiplicada por un peso  $w$ .

Cada neurona de la capa de entrada tiene un peso  $w$  asociado a cada una de las neuronas ocultas ( $w_{ih}$ )<sup>2</sup>, y cada neurona oculta tiene un peso asociado a la neurona de salida ( $w_{ho}$ ); estos pesos son clave ya que, son los que hacen que la red adquiera conocimiento.

<sup>2</sup> El subíndice  $i$  se refiere a la capa de entrada (input en inglés),  $h$  se refiere a la capa oculta (hidden en inglés) y el subíndice  $o$  se refiere a la capa de salida (output en inglés).

En el caso de la red neuronal de casos nuevos, se tienen 12 neuronas de entrada y 7 neuronas ocultas, por lo que en total se requieren 91 pesos  $w$  ( $12*7+7$ ). Por su parte en la red de estancia se tienen un total de 63 pesos  $w$  ( $8*12+7$ ). Estos pesos se inicializan con un valor aleatorio, y conforme la red evoluciona (aprende) adquieren valores positivos, negativos o cero.

Si  $w_{ih}$  es positivo, indica que la interacción entre las neuronas  $i$  y  $h$  es excitadora; es decir, siempre que la neurona  $i$  esté activada, la neurona  $h$  recibirá una señal de  $i$  que tenderá a activarla. Si  $w_{ih}$  es negativo, la sinapsis será inhibitoria. En este caso, si  $i$  está activada, enviará una señal a  $h$  que tenderá a desactivar a ésta. Finalmente, si  $w_{ih} = 0$ , se supone que no hay conexión entre ambas. Lo mismo sucede con los pesos  $w_{ho}$ .

**Cuadro 13: Pesos y entradas netas en la Red de casos nuevos**

Neuronas de entrada	Valor $x_i$	Pesos						
		$w_{ih1}$	$w_{ih2}$	$w_{ih3}$	$w_{ih4}$	$w_{ih5}$	$w_{ih6}$	$w_{ih7}$
Mes	0,005	-10,05	-4,93	40,67	1,55	58,28	24,44	-25,82
Promedio 3 años	0,527	3,15	-28,27	3,04	-21,53	29,73	-20,24	59,07
Promedio 2 año	0,527	25,28	20,01	15,13	-37,54	17,31	36,27	-3,85
Promedio 1 año	0,527	-12,83	-2,62	29,95	9,90	-40,34	-13,43	-6,37
Promedio 6 meses	0,527	-1,58	12,68	-7,23	21,84	-56,84	-17,96	16,05
Promedio 3 meses	0,509	11,16	4,58	25,71	-121,30	-8,40	0,31	42,13
Tasa general de mortalidad	0,004	148,85	19,50	5,11	20,91	-11,98	7,70	-24,17
Índice de Desarrollo Humano	0,00085	-5,70	-42,19	31,84	-26,72	97,09	37,37	1,94
Inversión Social en Salud	0,0048	1,64	-9,29	-31,65	-4,64	-8,76	0,10	10,79
Cobertura Régimen de Enfermedad y Maternidad	0,0862	-57,27	48,04	4,95	8,52	19,79	-23,87	7,26
Tasa de desempleo	0,0042	-13,11	-21,05	-6,47	-5,51	-23,60	46,61	-10,42
Nivel de pobreza	0,02	78,65	-61,20	239,78	1,40	15,03	26,31	18,13
<b>Valor (Net<sub>h</sub>)</b>		10,20	6,08	39,92	-75,36	-28,51	-9,09	56,42
<b><math>w_{ho}</math></b>		59,70	13,67	69,70	24,96	12,58	117,29	4,53
<b>Neuronas ocultas</b>		<b><math>O_{h1}</math></b>	<b><math>O_{h2}</math></b>	<b><math>O_{h3}</math></b>	<b><math>O_{h4}</math></b>	<b><math>O_{h5}</math></b>	<b><math>O_{h6}</math></b>	<b><math>O_{h7}</math></b>

Fuente: Los autores

Los valores de las neuronas de entrada se dividieron entre un factor común (entre 1000), ya que, si no se hacía esto durante el entrenamiento inicial de la red se generaban salidas

con valores sumamente altos, no reconocidos por el programa utilizado para hacer los cálculos (Microsoft Excel), y los pesos relativos adquirirían valores sumamente bajos.

Cada neurona en la capa oculta recibe señales de cada una de las neuronas de entrada. Las señales que llegan a cada neurona oculta se combinan entre ellas, generando así la entrada total,  $Net_h$ . ( $Net_h = \sum_i x_i w_{ih}$ ).

**Cuadro 14: Cálculo de la entrada neta para la neurona oculta  $O_{h1}$  en la Red de casos nuevos**

Valor de entrada ( $x_i$ )	$w_{ih1}$	$x_i * w_{ih1}$
0,005	-10,051	-0,050
0,527	3,149	1,659
0,527	25,275	13,320
0,527	-12,830	-6,762
0,527	-1,579	-0,832
0,509	11,161	5,681
0,004	148,853	0,595
0,00085	-5,702	-0,005
0,0048	1,643	0,008
0,0862	-57,267	-4,936
0,0042	-13,106	-0,055
0,02	78,650	1,573
$Net_{h1} (\sum_i x_i w_{ih})$		10,197

Fuente: Los autores

Al valor  $Net_h$  de cada neurona oculta se le aplica nuevamente una función de activación para enviar una señal a la neurona de salida. En este caso, para ambas redes, se utilizó la función identidad  $F(x) = x$ , por lo que  $Net_h$  es igual a  $y_h$

Multiplicando los valores de salida de las neuronas ocultas por los pesos relativos  $w_{ho}$  se generó la entrada total para la neurona de salida  $Net_o$  ( $Net_o = \sum_h y_h w_{ho}$ ).

A esta entrada total se le aplicó una función de salida  $F(Net_o)$  para generar el valor de salida final de la red (ya sea el número de casos nuevos de pacientes con cáncer hospitalizados en el mes, para la red de casos nuevos; o los días de hospitalización del

paciente con cáncer para la red de estancia). Nuevamente en ambas redes se utilizó la función identidad  $F(x) = x$ .

**Cuadro 15: Cálculo de una salida final en la Red de casos nuevos**

$Net_h (y_h)$	$w_{ho}$	$y_h * w_{ho}$
10,197	59,702	608,8
6,078	13,675	83,1
39,923	69,700	2782,6
-75,359	24,956	-1880,7
-28,508	12,582	-358,7
-9,090	117,294	-1066,2
56,417	4,528	255,5
Valor de la neurona salida ( $Net_o$ )		424,404

Fuente: Los autores

Hasta este punto se generó el primer valor pronosticado de la red. A partir de este momento inició un proceso de aprendizaje para lo cual se utiliza una regla determinada.

El proceso de aprendizaje consiste en hallar los pesos que codifican los conocimientos (disminuyen el error). Una regla de aprendizaje hace variar el valor de los pesos de una red hasta que estos adoptan un valor constante ( $dw/dt=0$ ), cuando esto ocurre se dice que la red ya ha aprendido (Sotolongo y Guzmán, 2001).

Para el caso de las redes neuronales de casos nuevos y de estancia se utilizó el modelo de aprendizaje conocido como propagación hacia atrás o back propagation.

#### 4.3.4 Entrenamiento de la red

En el modelo de propagación hacia atrás todas las neuronas interfieren en una disminución del error, por medio de la modificación de sus pesos; de ahí que el trabajo es dividido en muchas pequeñas tareas que cada neurona realiza.

Para aplicar el algoritmo back-propagation primero se compara el resultado obtenido en las neuronas de salida con la salida real de la red y se calcula un valor del error para cada neurona oculta y de salida.

Para las neuronas de la capa de salida, el valor del error (delta) es:

$$\delta_o = (d_o - y_o)f'_o(Net_o)$$

donde  $d_o$  es el valor real y  $y_o$  es el valor de salida de la red ( $y_o = Net_o$ ).

Continuando con el ejemplo de la sección 1.3.3, la salida real era de 515 casos nuevos y la calculada por la red fue de 424 casos nuevos.

Tanto para la red de estancia como para la red de casos nuevos, la derivada de la función de salida  $f'_o(Net_o)$  es igual a 1, ya que,  $f$  es la función identidad y su derivada es 1.

Por lo tanto para el ejemplo el error de la red sería de 91 casos (515-424).

Para las neuronas ocultas, el error (delta) está dado por:

$$\delta_h = f'_h(Net_h) \sum \delta_o w_{ho}$$

Donde se observa que el error en las capas ocultas depende de todos los términos de error de la capa de salida. De aquí surge el término propagación hacia atrás.

Nuevamente, en el caso de las redes de casos nuevos y de estancia,  $f'_h(Net_h)$  es igual a 1, ya que, se utilizó la función identidad por lo que el error ( $\delta_h$ ) de cada neurona oculta se calculó como se presenta en el siguiente cuadro.

**Cuadro 16: Cálculo del error para las neuronas ocultas Red de casos nuevos**

Neurona oculta	$\delta_o$	$w_{ho}$	$\delta_h (\delta_o * w_{ho})$
O <sub>h1</sub>	90,596	59,702	5408,710
O <sub>h2</sub>	90,596	13,675	1238,850
O <sub>h3</sub>	90,596	69,700	6314,481
O <sub>h4</sub>	90,596	24,956	2260,909
O <sub>h5</sub>	90,596	12,582	1139,886
O <sub>h6</sub>	90,596	117,294	10626,291
O <sub>h7</sub>	90,596	4,528	410,232

Fuente: Los autores

Posteriormente, estos errores se transmiten hacia atrás por medio de un ajuste en los pesos de la red, partiendo de la capa de salida, hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida, recibiendo el porcentaje de error aproximado a la participación de la neurona intermedia en la salida original.

Los pesos se ajustan de la forma siguiente:

Para los pesos de las neuronas de la capa oculta:

$$w_{ho}(t + 1) = w_{ho}(t) + \Delta w_{ho}(t + 1)$$

$$\Delta w_{ho}(t + 1) = \alpha \delta_o y_h$$

donde  $\delta_o$  es el error de salida,  $y_h$  es el valor de salida de las neuronas (que es igual a  $Net_h$ ) y  $\alpha$  es una constante que corresponde a la tasa de aprendizaje, es decir a la velocidad con que se espera que la red neuronal aprenda.

En el caso de las redes neuronales de estancia y de casos nuevos, al inicio se utilizó un valor aleatorio para la constante  $\alpha$  entre 0 y 1, pero, se determinó que se generan mejores resultados iniciando con un aleatorio de orden de  $10^{-5}$ .

Continuando con el ejemplo, utilizando un  $\alpha$  de  $2,81 \times 10^{-5}$  y considerando que  $\delta_o$  es igual 90, el valor del diferencial  $\Delta w_{ho}(t + 1)$  para las neuronas ocultas era el siguiente:

**Cuadro 17: Cálculo del error para las neuronas ocultas Red de casos nuevos**

$fa$	$y_h$ ó $Net_h$	$\delta_o$	$\Delta w_{ho}$
2.81E-05	10.197	90.596	0.0259
2.81E-05	6.078	90.596	0.0155
2.81E-05	39.923	90.596	0.1015
2.81E-05	-75.359	90.596	-0.1916
2.81E-05	-28.508	90.596	-0.0725
2.81E-05	-9.090	90.596	-0.0231
2.81E-05	56.417	90.596	0.1434

Fuente: Los autores

El proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total.

Para los pesos de las neuronas de la capa de entrada, el error (delta) estaría dado por:

$$w_{ih}(t + 1) = w_{ih}(t) + \Delta w_{ih}(t + 1);$$

$$\Delta w_{ih}(t + 1) = \alpha \delta_h x_i$$

En los siguientes cuadros se presentan los cálculos obtenidos para el ejemplo de la red de casos nuevos.

Cuadro 18: Cálculo del error para las neuronas de entrada en la Red de casos nuevos

Neuronas entrada	Valor $x_i$	Diferencial de los pesos de las variables de entrada $\Delta w_{ih}$						
		$\Delta w_{ih1}$	$\Delta w_{ih2}$	$\Delta w_{ih3}$	$\Delta w_{ih4}$	$\Delta w_{ih5}$	$\Delta w_{ih6}$	$\Delta w_{ih7}$
Mes	0.005	0.000759	0.000174	0.000886	0.000317	0.000160	0.001491	0.000058
Prom 3 a	0.527	0.079985	0.018320	0.093379	0.033435	0.016857	0.157143	0.006067
Prom 2 a	0.527	0.079985	0.018320	0.093379	0.033435	0.016857	0.157143	0.006067
Prom 1 a	0.527	0.079985	0.018320	0.093379	0.033435	0.016857	0.157143	0.006067
Prom 6 m	0.527	0.079985	0.018320	0.093379	0.033435	0.016857	0.157143	0.006067
Prom 3 m	0.509	0.077253	0.017695	0.090190	0.032293	0.016281	0.151776	0.005859
TM	0.004	0.000607	0.000139	0.000709	0.000254	0.000128	0.001193	0.000046
IDH	0.00085	0.000129	0.000030	0.000151	0.000054	0.000027	0.000253	0.000010
ISS	0.0048	0.000729	0.000167	0.000851	0.000305	0.000154	0.001431	0.000055
CREM	0.0862	0.013083	0.002997	0.015274	0.005469	0.002757	0.025703	0.000992
TD	0.0042	0.000637	0.000146	0.000744	0.000266	0.000134	0.001252	0.000048
Np	0.02	0.003035	0.000695	0.003544	0.001269	0.000640	0.005964	0.000230
$fa$		0.0000281	0.0000281	0.0000281	0.0000281	0.0000281	0.0000281	0.0000281
$\delta_{ho}$		5408.71	1238.85	6314.48	2260.91	1139.89	10626.29	410.23
Neuronas ocultas		$O_{h1}$	$O_{h2}$	$O_{h3}$	$O_{h4}$	$O_{h5}$	$O_{h6}$	$O_{h7}$

Fuente: Los autores

Cuadro 19: Cálculo del  $\Delta w_{ih}$  para la neurona de entrada 2 (Prom 3a) en la Red de casos nuevos

	$fa$	$\delta_{ho}$	$x_i$	$\Delta w_{ih}$
$O_{h1}$	0.0000281	5408.71	0.527	0.07998
$O_{h2}$	0.0000281	1238.85	0.527	0.01832
$O_{h3}$	0.0000281	6314.48	0.527	0.09338
$O_{h4}$	0.0000281	2260.91	0.527	0.03343
$O_{h5}$	0.0000281	1139.89	0.527	0.01686
$O_{h6}$	0.0000281	10626.29	0.527	0.15714
$O_{h7}$	0.0000281	410.23	0.527	0.00607

Fuente: Los autores

Finalmente, para determinar los pesos a utilizar en la siguiente corrida de la red, se le suma el diferencial de cada neurona de entrada y cada neurona oculta al peso anterior:

$$w(t + 1) = w(t) + \Delta w$$

El proceso se repite hasta que el término de error resulta aceptablemente pequeño para cada uno de los patrones aprendidos.

$$E_p = \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2$$

El proceso de aprendizaje descrito se realizó con los datos de 168 meses disponibles para la Red de casos nuevos de pacientes con cáncer.

En el caso de la Red de estancia, aunque se disponía información sobre 71631 casos de los últimos 14 años, se trabajó únicamente con los datos de pacientes egresados en el 2010 ya que, a diferencia de la otra red, la variable tiempo no tiene incidencia en el fenómeno de estudio. En este caso podrían afectar otras variables relacionadas con proceso, gestión interna o las características propias del paciente, por lo que al usar datos del 2010 se busca que el modelo represente la realidad actual de los hospitales y de la salud de la población en general.

Se seleccionaron los últimos 1000 casos del 2010, ya que, además de la indicación realizada en el párrafo anterior, el programa informático utilizado para la creación de la red (Microsoft Excel) tiene una capacidad limitada para manejar grandes volúmenes de datos relacionados con múltiples fórmulas matemáticas y macros.

#### **4.3.5 Evaluación inicial de la red**

Para realizar la evaluación de los modelos de redes neuronales desarrollados se utilizaron los siguientes tipos de error:

- Error Porcentual Medio Absoluto (MAPE), el cual permite establecer la diferencia media entre el pronóstico y el valor real en términos porcentuales.
- Desviación Media Absoluta (MAD), con el cual se determina la diferencia media en unidades entre el pronóstico y el valor real.

- Error Medio Cuadrado (MSE), el cual muestra la sumatoria de los errores de pronóstico al cuadrado.

Primero se evalúan los resultados generados durante la construcción y luego los generados después del entrenamiento de la red.

Para el caso de la red de casos nuevos se evaluó la construcción del modelo con los datos de los primeros 120 meses (equivalente al 71% de los disponibles). El entrenamiento se evaluó con los datos de los siguientes 36 meses (21%). Los restantes 12 períodos (8%) se utilizaron posteriormente para comparar las proyecciones con los valores reales.

En el caso de la red de estancia, la construcción se evaluó con datos de 700 pacientes hospitalizados (70%) y el entrenamiento se evaluó con 150 casos (15%). El 15% restante se utilizó posteriormente para evaluar la efectividad de las proyecciones.

En el siguiente cuadro se presentan los resultados de la evaluación inicial de la red, antes de optimizarla, para las dos redes neuronales desarrolladas.

**Cuadro 20: Evaluación inicial de las redes de casos nuevos y de estancia hospitalaria**

Red	Tipo de error	Error durante la construcción del modelo	Error después del entrenamiento
Red casos nuevos	MAPE	11.76%	7.53%
	MAD	53.68	41.53
	MSE	18618	2515
Red estancia	MAPE	84.13%	47.37%
	MAD	2.18	1.09
	MSE	8.66	2.22

Fuente: Los autores

Cómo se observa en el cuadro anterior, ambas redes disminuyeron considerablemente el error después del entrenamiento.

En el caso de la red de casos nuevos, el error absoluto porcentual de la media (MAPE) pasó de 11,76% durante la construcción del modelo a 7.53% después del entrenamiento. El valor del MAD disminuyó de 53.68 a 41.53 puntos y el del MSE pasó de 18618 a 2515.

Por otro parte, para la red de estancia el MAPE pasó de 84% durante la construcción a 47% después del entrenamiento. El MAD pasó de 2.18 a 1.09 y el MSE de 8.66 a 2.22. A pesar de que un 47% de error sigue siendo considerable, el hecho de que disminuyera 37% indica que el aprendizaje de la red fue efectivo.

Una hipótesis de que el porcentaje de error de la red de estancia sea alto, es la cantidad de casos utilizados en el aprendizaje (1000). Se esperaría que entre más casos se incluyan en el proceso de construcción y aprendizaje, dicho porcentaje de error disminuiría.

#### **4.4 Optimización de la red mediante algoritmos genéticos (AG)**

El uso de la técnica de algoritmos genéticos está más generalizado en problemas de optimización, ya sea para maximizar beneficios o minimizar errores. Su aplicación implica la definición de conjuntos de solución del problema y de una función de evaluación que permita discriminar entre las diferentes soluciones planteadas.

Esta técnica de simple aplicación pero con resultados robustos, probados en diferentes situaciones, se utiliza para optimizar los resultados de las dos redes neuronales objeto de esta investigación, básicamente para buscar un conjunto solución inicial que permita a la red generar resultados satisfactorios en un tiempo más corto, o sea, acelerando la adaptación de la red a las condiciones de las diferentes variables que influyen en el resultado esperado.

Para la aplicación de esta técnica se tienen que definir una serie de elementos, los cuales conforman el algoritmo, tales como la población inicial, la función de evaluación, los procedimientos de selección de individuos, las condiciones de los cruces y la mutación,

que van a permitir al algoritmo moverse a través del espacio de solución, donde se pueda encontrar la solución óptima.

El algoritmo definido tanto para la red de casos nuevos como para la de estancia hospitalaria, fue un algoritmo de minimización del error de pronóstico o estimación.

Es importante destacar que para la elaboración del algoritmo genético las soluciones son una combinación de los valores de los diferentes pesos que relacionan las variables de entrada con las neuronas ocultas ( $w_{ih}$ ) y a su vez estas con la capa de salida ( $w_{ho}$ ), e incluyendo la tasa de aprendizaje ( $fa$ ), ya que, en las pruebas realizadas con las redes neuronales se determinó que esta tasa de aprendizaje influía de manera importante en los resultados.

Algo necesario de resaltar es que la optimización se hizo con base en toda la red, es decir, incluyendo tanto la etapa de construcción como la etapa de entrenamiento, para que no se perdiera la consistencia de la red de aprender de todos los casos históricos y con ello generara una mejor adaptación.

#### **4.4.1 Tamaño de población**

El tamaño de la población es la cantidad de soluciones (conjunto de pesos  $w$  y  $fa$ ) que se van a estar analizando en cada etapa del algoritmo. Para ambas redes el tamaño de población se definió en ocho individuos, de manera que se analizara un importante número de soluciones, pero, sin hacer demasiado lento y costoso el proceso.

#### **4.4.2 Población inicial**

La población inicial es el conjunto de soluciones sobre el cual se inicia el algoritmo genético, en el caso de la red de casos nuevos, la población inicial se definió con números aleatorios con valores entre -50 y 100 para los pesos que relacionen las diferentes neuronas, y un aleatorio entre 0 y 1, multiplicado por  $10^{-4}$  para la tasa de aprendizaje.

En el caso de la red de estancia, la población inicial se definió con valores entre -2 y 2 para los pesos de las relaciones entre las neuronas y un aleatorio entre 0 y 1 multiplicado por  $10^{-3}$  para la tasa de aprendizaje.

Esto valores fueron obtenidos analizando los resultados del aprendizaje que se había realizado para ambas redes, tratando de utilizar rangos de posibles cantidades que incluyeran los valores observados en las variables posterior al entrenamiento efectuado a la red.

Para ilustrar mejor este proceso se desarrolla un ejemplo de la optimización de la red de casos nuevos, pero, ejemplificado con solo una parte de los pesos que se deben ir modificando. La población inicial sería la siguiente:

**Cuadro 21: Definición de la población inicial**

Cromosoma	FA	W1	W2	W3	W4	W5
1	1,7329E-05	-6,55979	-19,594	-9,7086	-46,472	9,8176
2	2,1713E-05	14,9468	-3,37171	-46,188	3,451	13,756
3	3,6472E-05	-29,2447	40,4086	-47,805	46,352	-34,428
4	2,3098E-05	-20,9298	-19,0834	49,8861	-15,016	-45,583
5	4,0554E-05	7,07829	35,8608	-9,2617	-33,74	-19,701
6	1,4175E-05	-37,1374	27,3395	-35,429	10,617	-6,686
7	5,6424E-06	-15,5153	29,6089	-30,053	-41,071	39,307
8	8,3497E-05	46,811	29,0125	31,2843	-21,942	33,752

Fuente: Los autores

#### 4.4.3 Evaluación de individuos

Para realizar la evaluación de los individuos se utiliza la función objetivo o función de evaluación, la cual para ambos modelos es la función que define el error de pronóstico conocido como Error Porcentual Medio Absoluto o MAPE (por si siglas en inglés).

$$MAPE = \frac{\sum |PE|}{n}$$

Para aplicar esta función se utilizó el error que se presenta en los datos correspondientes al entrenamiento de cada red, o sea el último 20% en la red de casos nuevos y 15% en el caso de la red de estancia.

El valor se obtuvo definiendo cada una de las soluciones como el conjunto inicial de pesos de la red neuronal, con lo cual se generaba el error del entrenamiento, que es el valor que se deseaba minimizar.

En el ejemplo los datos serían los siguientes:

**Cuadro 22: Evaluación de la población inicial**

Cromosoma	Error	FA	W1	W2	W3	W4	W5
1	28,94%	1,7329E-05	-6,55979	-19,594	-9,7086	-46,472	9,8176
2	9,63%	2,1713E-05	14,9468	-3,37171	-46,188	3,451	13,756
3	11,48%	3,6472E-05	-29,2447	40,4086	-47,805	46,352	-34,428
4	18,53%	2,3098E-05	-20,9298	-19,0834	49,8861	-15,016	-45,583
5	17,39%	4,0554E-05	7,07829	35,8608	-9,2617	-33,74	-19,701
6	13,67%	1,4175E-05	-37,1374	27,3395	-35,429	10,617	-6,686
7	110,85%	5,6424E-06	-15,5153	29,6089	-30,053	-41,071	39,307
8	22,13%	8,3497E-05	46,811	29,0125	31,2843	-21,942	33,752

Fuente: Los autores

#### 4.4.4 Selección

La selección de los individuos se realiza de acuerdo con la evaluación de los mismos, y dada la naturaleza de la función de evaluación, lo que se busca es que los individuos minimicen el error de pronóstico, por lo tanto se seleccionan aquellos que generan los errores menores, de manera que estos continúen en el proceso. Para efectos de ambas redes se definió la selección del 50% de los individuos de la población para continuar en el proceso y para que funcionaran como padres al momento de realizar los cruces. Esto permitió que el algoritmo genético en esta etapa renovara la mitad de los individuos de la población en estudio, creando un balance entre lo que se tiene y lo nuevo por explorar.

Para el ejemplo que se está desarrollando, se necesitaría ordenar la población de acuerdo con el error y seleccionar los individuos con mejores resultados, tal como se muestra en el cuadro 23.

**Cuadro 23: Selección de los individuos**

Cromosoma	Error	FA	W1	W2	W3	W4	W5
2	9,63%	2,1713E-05	14,9468	-3,37171	-46,188	3,451	13,756
3	11,48%	3,6472E-05	-29,2447	40,4086	-47,805	46,352	-34,428
6	13,67%	1,4175E-05	-37,1374	27,3395	-35,429	10,617	-6,686
5	17,39%	4,0554E-05	7,07829	35,8608	-9,2617	-33,74	-19,701
4	18,53%	2,3098E-05	-20,9298	-19,0834	49,8861	-15,016	-45,583
8	22,13%	8,3497E-05	46,811	29,0125	31,2843	-21,942	33,752
1	28,94%	1,7329E-05	-6,55979	-19,594	-9,7086	-46,472	9,8176
7	110,85%	5,6424E-06	-15,5153	29,6089	-30,053	-41,071	39,307

Fuente: Los autores

#### 4.4.5 Recombinación o cruce

El cruce es una combinación de los individuos con mejores resultados (padres) que busca que se generen nuevos individuos (hijos) con mejores resultados que aquellos que les dieron origen.

Para ambas redes se realizaron dos cruces, donde se generaron cuatro nuevos individuos, de manera que se volviera a completar el tamaño de población a ocho.

En cuanto a la forma del cruce, se especificó una función de combinación definida como una proporción entre los valores de cada uno de los padres, para cada uno de los elementos que conforman el individuo, la proporción estuvo definida por la fórmula:

$$H_1 = P + \alpha(P - M)$$

$$H_2 = P - \alpha(P - M)$$

Donde:

$H_1$  = Valor del elemento del primer hijo.

H<sub>2</sub> = Valor del elemento del segundo hijo.

P = Valor del elemento del padre.

M = Valor del elemento de la madre.

$\alpha$  = Número aleatorio entre 0 y 1.

Para efectos de la escogencia de individuos que funcionaran como padre o madre se definió que de los cuatro mejores resultados, que se seleccionaban para continuar en el proceso, el mejor de ellos iba a funcionar como padre en ambos cruces; mientras que en el primer cruce el segundo mejor resultado funcionaba de madre y para el segundo cruce, la madre era el tercer mejor resultado. Esta combinación permitía darle mayor énfasis al mejor resultado encontrado, buscando nuevos espacios de solución en la zona cercana a este, que trataran de converger más rápidamente a resultados mejores.

De estos dos cruces surgen cuatro nuevos individuos, los cuales pasan a formar parte de la población de estudio y se evalúan en la red neuronal, para determinar el error que puedan proporcionar. Además, estos individuos son ordenados según el error generado con tal de diferenciar los mejores individuos que se van obteniendo.

Para el ejemplo, los datos de los nuevos individuos, debidamente evaluados, se muestran en el cuadro 24.

**Cuadro 24: Evaluación de los nuevos individuos**

Cromosoma	Error	FA	W1	W2	W3	W4	W5
2	9,63%	2,1713E-05	14,9468	-3,37171	-46,188	3,451	13,756
3	11,48%	3,6472E-05	-29,2447	40,4086	-47,805	46,352	-34,428
6	13,67%	1,4175E-05	-37,1374	27,3395	-35,429	10,617	-6,686
5	17,39%	4,0554E-05	7,07829	35,8608	-9,2617	-33,74	-19,701
9	16,50%	1,2793E-05	41,6544	-29,8308	-45,211	-22,477	42,877
10	13,88%	3,0633E-05	-11,7608	23,0874	-47,165	29,379	-15,364
11	12,41%	2,3541E-05	27,5759	-10,8184	-48,797	1,7134	18,713
12	8,60%	1,9885E-05	2,31769	4,07499	-43,579	5,1886	8,7996

Fuente: Los autores

#### 4.4.6 Mutación

Para efectuar el proceso de mutación en ambas redes se utilizó una tasa de mutación del 33%, la cual responde a la necesidad de equilibrar entre la ampliación de las posibilidades de encontrar mejores soluciones en otras áreas del conjunto solución, con el control de la eficiencia del proceso, de manera que este no se haga muy lento y costoso.

Para efectos de definir los individuos que se iban a mutar, se respetó los dos mejores individuos que se tenían en la población de estudio, los cuales no se mutaron, y se aplicó la mutación a los demás individuos.

Para definir los individuos y los pesos exactos que se iban a estar mutando, se aplicó un proceso de escogencia aleatoria; adicionalmente el valor a mutar se sustituyó con un valor aleatorio tal como los definidos en la población inicial. El detalle del ejemplo se observa en el siguiente cuadro.

**Cuadro 25: Detalle de la mutación de los individuos**

Cromosoma	FA	W1	W2	W3	W4	W5
12	1,9885E-05	2,31769	4,07499	-43,579	5,1886	8,7996
2	2,1713E-05	14,9468	-3,37171	-46,188	3,451	13,756
3	3,6472E-05	-29,2447	40,4086	-47,805	46,352	-34,428
11	2,1488E-05	27,5759	-10,8184	-48,797	1,7134	18,713
6	1,4175E-05	-42,5268	-6,44925	-35,429	10,617	31,335
10	3,0633E-05	-11,7608	23,0874	-47,165	39,7	49,102
9	1,2793E-05	41,6544	-49,3113	-0,4376	-22,477	42,877
5	4,0554E-05	7,07829	35,8608	-9,2617	-12,281	-19,701

Fuente: Los autores

#### 4.4.7 Evaluación de nueva generación

Una vez que se ha completado la mutación se vuelve a evaluar los seis individuos que fueron mutados en la red neuronal correspondiente, para determinar si alguno de ellos genera porcentajes de error menores a los dos mejores resultados que se tenía y se mantuvieron sin cambio.

La evaluación de la nueva generación se observa en el cuadro 26.

**Cuadro 26: Evaluación de la nueva generación**

Cromosoma	Error	FA	W1	W2	W3	W4	W5
12	8,60%	1,9885E-05	2,31769	4,07499	-43,579	5,1886	8,7996
2	9,63%	2,1713E-05	14,9468	-3,37171	-46,188	3,451	13,756
6	11,13%	1,4175E-05	-42,5268	-6,44925	-35,429	10,617	31,335
5	16,23%	4,0554E-05	7,07829	35,8608	-9,2617	-12,281	-19,701
9	22,01%	1,2793E-05	41,6544	-49,3113	-0,4376	-22,477	42,877
10	22,43%	3,0633E-05	-11,7608	23,0874	-47,165	39,7	49,102
3	24,95%	3,6472E-05	-29,2447	40,4086	-47,805	46,352	-34,428
11	28,74%	2,1488E-05	27,5759	-10,8184	-48,797	1,7134	18,713

Fuente: Los autores

#### 4.4.8 Evaluación de la red neuronal optimizada

La secuencia de tareas descrita anteriormente de selección, cruce, evaluación, mutación y reevaluación, constituye una iteración del algoritmo, esta secuencia se debe seguir ejecutando de forma continua, de manera que cada vez se vayan obteniendo mejores resultados, hasta un punto en donde ya no se disminuya el error o se esté satisfecho con el resultado obtenido.

En el caso de las dos redes neuronales objeto de esta investigación, se aplicaron bloques de 45 iteraciones para evaluar hasta donde se optimizaba los resultados, siempre que estos procesos no implicaran un consumo excesivo de tiempo y esfuerzo que no estuvieran presentando una mejora significativa en los resultados.

En el caso de la red de casos nuevos el resultado obtenido después del proceso de aprendizaje de la red, fue de un error de pronóstico (MAPE) de 7.53%; sin embargo, al aplicar un primer bloque de 45 iteraciones el resultado mejoró hasta valores cercanos al 6.62%, posteriormente al aplicarle 45 iteraciones más, se llegó a un valor de 5.62%, luego 5.46%, 5.21% y así sucesivamente se siguieron aplicando iteraciones hasta mejorar los resultados, que en este caso específico después de varios procesos se logró llegar hasta un

porcentaje de error en el entrenamiento de 5.07%, que equivale a un error global de la red de 7.02%.

Por otra parte, para la red de estancia se arrancó con los valores de aprendizaje que eran de 47,37%, posteriormente se le aplicó un proceso de optimización de 45 iteraciones con el cual se logró obtener un porcentaje de error de 44,91%, en este caso al seguir aplicando 45 iteraciones se logró un porcentaje de 41,87% y así sucesivamente al aplicar iteraciones se logró llegar a un porcentaje mínimo de error MAPE de 40,79%.

El análisis anterior se presenta en el siguiente cuadro.

**Cuadro 27: Evaluación inicial de las redes de casos nuevos y de estancia hospitalaria**

Red	Error antes de la optimización (MAPE)	Error después de la optimización (MAPE)
Red casos nuevos	7.53%	5.07%
Red estancia	47.37%	40.79%

Fuente: Los autores

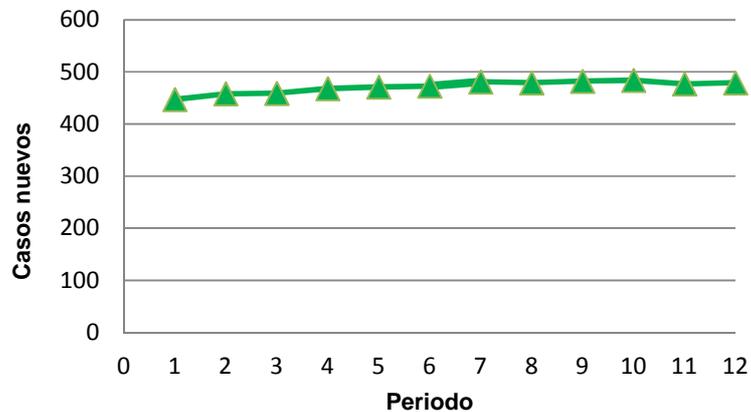
#### 4.4.9 Generación de pronósticos mediante el modelo de redes neuronales

En el siguiente cuadro y gráfico se presentan las proyecciones obtenidas para 12 períodos, utilizando el modelo de redes neuronales con datos de los primeros 156 meses.

**Cuadro 28: Proyecciones de casos nuevos de pacientes hospitalizados con cáncer utilizando para 12 períodos de análisis (año 2011).**

Período	Casos nuevos
1	447
2	458
3	459
4	468
5	471
6	473
7	481
8	479
9	482
10	484
11	477
12	479

**Gráfico 22: Proyecciones de casos nuevos de pacientes hospitalizados con cáncer para 12 períodos de análisis (año 2011).**



Fuente: Los autores

Como se observa en el cuadro 28 y en el gráfico 22, en los pronósticos se presenta una leve tendencia ascendente en la cantidad de casos nuevos por mes con un pico en el período 7, pero, con valores que se concentran entre los 450 y los 500 casos nuevos por mes para el año pronosticado.

El error (MAPE) de pronóstico, considerando los valores reales, fue de 5.64%, el MAD de 23.3 y el MSE de 1182.8.

Es importante mencionar que para generar proyecciones mayores a un mes, en un mismo momento, utilizando el modelo de redes neuronales optimizadas con algoritmos genéticos, debe considerarse que las variables de entrada, relacionadas con el comportamiento de los datos, deben irse actualizando con los valores pronosticados en el período anterior, es decir que al generar el pronóstico del segundo mes, los promedios van a correrse hasta incluir el primer mes pronosticado y así sucesivamente, no obstante, si se hacen los pronósticos un mes a la vez, se pueden utilizar los datos reales, de manera que se mejore el pronóstico.

En el caso de la red de estancia, con el modelo se generaron proyecciones determinando los días de hospitalización de 150 pacientes, los cuales se compararon con los valores reales. Los errores de pronóstico fueron los siguientes: MAPE: 26,84%; MAD: 0,55 y MSE: 0,47. Como se observa en estos resultados, los errores de pronóstico, fueron mejores que los obtenidos antes y después del entrenamiento de la red de estancia.

#### **4.5 Comparación de resultados de la red neuronal con los generados por un programa informático especializado en pronósticos**

Para analizar la efectividad del modelo de red neuronal de casos nuevos de pacientes con cáncer, se compararon los resultados del mismo con los obtenidos mediante 4 modelos tradicionales de pronósticos descritos en el marco teórico de esta investigación: Suavización Exponencial Simple, Suavización Exponencial Doble, Winters Aditivo y Winters Multiplicativo.

Para generar los pronósticos con estos 4 modelos tradicionales se utilizó el programa estadístico MINITAB versión 15.0, el cual cuenta con un módulo especializado en pronósticos.

En todos los casos se pronostico con base en los datos de 156 meses, y no con los 168 disponibles, esto con el fin de poder proyectar 12 períodos y compararlos con los resultados reales de casos nuevos de pacientes con cáncer.

Para realizar la comparación se consideraron 3 tipos de errores:

- Error absoluto porcentual de la media (MAPE)
- Error absoluto de la media (MAD)
- Error cuadrático de la media (MSE)

En el anexo 1 de este documento se presentan los resultados generados con Minitab, incluyendo el pronóstico, los errores y el gráfico respectivo, para cada uno de los 4 modelos tradicionales indicados.

En el siguiente cuadro y gráfico se presentan los resultados de los pronósticos de los cuatro modelos tradicionales junto con el de la red neuronal.

**Cuadro 29: Comparación de resultados de pronósticos generados por los modelos tradicionales y el modelo de redes neuronales**

Período	1	2	3	4	5	6	7	8	9	10	11	12	Errores totales		
													MAPE	MAD	MSE
SES	481	481	481	481	481	481	481	481	481	481	481	481	7.53	32.42	1758
SED	510	511	513	514	515	516	517	518	519	520	521	523	9.57	41.19	2806
WM	440	497	549	493	541	538	535	524	539	554	573	556	6.34	27.93	1253
WA	447	498	545	498	540	537	535	527	541	555	570	557	6.24	27.39	1207
<b>RN</b>	<b>447</b>	<b>458</b>	<b>459</b>	<b>468</b>	<b>471</b>	<b>473</b>	<b>481</b>	<b>479</b>	<b>482</b>	<b>484</b>	<b>477</b>	<b>479</b>	<b>5.07</b>	<b>22.9</b>	<b>818</b>
Real	351	422	492	458	477	434	470	483	460	483	472	496			

Fuente: Los autores

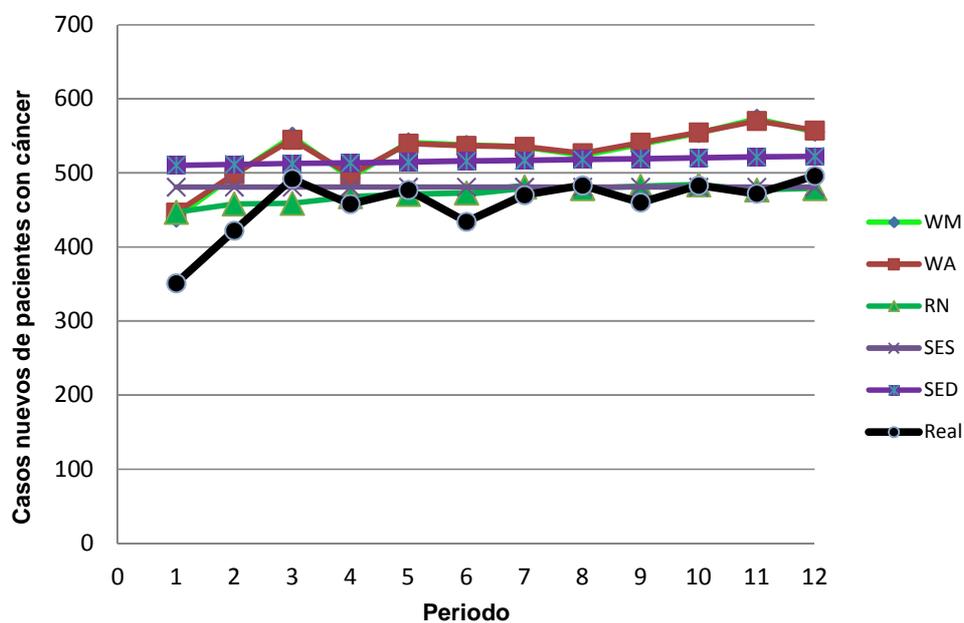
Cómo se observa en el cuadro 29, el modelo de redes neuronales generó los mejores resultados en términos del error de pronóstico (MAPE 5.07, MAD 22.9 y MSE 818). Posteriormente se encuentra el método de Winters Aditivo (MAPE 6.24, MAD 27.39 y MSE 1207), luego le sigue el método Winters Multiplicativo (MAPE 6.34, MAD 27.93 y MSE 1253), después el de Suavización Exponencial Simple (MAPE 7.53, MAD 32.42 y MSE 1758) y por último el de Suavización Exponencial Doble (MAPE 9.57, MAD 41.19 y MSE 2806).

De acuerdo con estos resultados, puede concluirse que el modelo de red neuronal optimizada con algoritmos genéticos es más efectivo para la proyección de casos nuevos de pacientes con cáncer, ya que, presenta un menor error de pronóstico.

Con esto se confirma la hipótesis nula planteada en la investigación, no obstante, se debe hacer la aclaración de que la evaluación del modelo se hace con base en indicadores relativos como el promedio de los errores de pronóstico y no utilizando una prueba estadística de hipótesis específica.

Para probar la hipótesis también se utilizó el método intuitivo por medio de la generación y análisis del siguiente gráfico, en el que se observa como el comportamiento de los pronósticos del modelo de redes neuronales es el que se acercó más a los valores de los datos reales.

**Gráfico 23 : Comparación de resultados de pronósticos generados por los modelos tradicionales y el modelo de redes neuronales**



Fuente: Los autores

Para sustentar la conclusión indicada en el párrafo anterior, en el siguiente cuadro se presentan los errores del pronóstico real, es decir, los errores resultantes de comparar la cantidad de casos nuevos de cáncer pronosticados en cada modelo con los casos reales, considerando los últimos 12 períodos disponibles.

**Cuadro 30: Comparación de errores de pronósticos generados por los modelos tradicionales y el modelo de redes neuronales**

Modelo	Período												MAPE
	1	2	3	4	5	6	7	8	9	10	11	12	
SES	37%	14%	2%	5%	1%	11%	2%	0%	5%	0%	2%	3%	6.9%
SED	45%	21%	4%	12%	8%	19%	10%	7%	13%	8%	10%	5%	13.6%
WM	25%	18%	12%	8%	13%	24%	14%	8%	17%	15%	21%	12%	15.6%
WA	27%	18%	11%	9%	13%	24%	14%	9%	18%	15%	21%	12%	15.8%
RN	27%	9%	7%	2%	1%	9%	2%	1%	5%	0%	1%	3%	5.64%

													MAD
SES	129.8	58.8	11.2	22.8	3.8	46.8	10.8	2.2	20.8	2.2	8.8	15.2	27.8
SED	159.3	89.4	20.5	55.6	37.7	81.9	47.0	35.1	59.2	37.3	49.4	26.6	58.3
WM	88.6	75.2	57.2	35.3	64.3	103.7	65.0	40.8	78.8	71.0	101.2	59.6	70.1
WA	95.6	76.3	52.8	39.5	62.7	102.6	65.2	43.6	80.8	71.6	98.4	61.4	70.9
RN	96.0	36.0	33.0	10.0	6.0	39.0	11.0	4.0	22.0	1.0	5.0	17.0	23.3

													MSE
SES	16845	3456	126	519	14	2189	116	5	432	5	77	231	2001,3
SED	25373	7993	421	3095	1425	6702	2207	1232	3506	1393	2444	705	4708,0
WM	7846	5657	3273	1246	4135	10750	4219	1661	6216	5040	10246	3548	5319,8
WA	9137	5815	2790	1563	3929	10533	4246	1904	6528	5132	9677	3771	5418,8
RN	9216	1296	1089	100	36	1521	121	16	484	1	25	289	1182,8

Fuente: Los autores

Cómo se observa, para los tres tipos de error calculados, la red neuronal de casos nuevos presentó mejores resultados (MAPE 5.64%, MAD 23.3 y MSE 1182.8) que los otros 4 modelos tradicionales analizados. Después de la red neuronal, el orden de los modelos tradicionales según el error presentado fue el siguiente: Suavización Exponencial Simple, Suavización Exponencial Doble, Winters Multiplicativo y Winters Aditivo.

En este punto es importante mencionar, que si bien es cierto se determinó que utilizando los algoritmos genéticos es posible optimizar los pesos iniciales para las conexiones entre neuronas de la red, de modo que el error total de la red, antes y después del entrenamiento se disminuyera, ayudando a que los pronósticos que se generaron mediante el modelo de redes neuronales presentaran un menor valor de error; la

optimización tuvo un efecto en la representación del fenómeno de estudio, ya que, tendió a disminuir la variabilidad u oscilación de los datos, y aunque en términos generales el error disminuyó, no se logró representar fielmente los “picos” que se presentan en algunos meses.

Por otra parte, si la red no se optimiza, es decir si únicamente se entrena con back-propagation, se logran representar “los picos” y el comportamiento de los datos del fenómeno en estudio, pero con una menor exactitud, lo que incrementa el error de pronóstico.

#### 4.5.1 Evaluación del horizonte del pronóstico

Para evaluar el horizonte de pronóstico que se recomienda utilizar con el modelo de redes neuronales desarrollado, se generaron pronósticos tomando como datos de entrada al modelo los primeros 132 meses, con el fin de utilizar los restantes 36 disponibles para hacer la comparación y evaluación de lo pronosticado con lo real en tres escenarios: un año, dos años, y tres años (cada año tiene 12 períodos). Para generar los pronósticos de los modelos tradicionales se utilizó el programa informático Minitab.

En el siguiente cuadro se presentan los resultados del error de pronósticos para cada modelo en los diferentes horizontes.

**Cuadro 31: Evaluación del horizonte de pronósticos del modelo de redes neuronales en comparación con los modelos tradicionales**

Modelo	Un año			Dos años			Tres años		
	MAPE	MAD	MSE	MAPE	MAD	MSE	MAPE	MAD	MSE
SES	6,9%	27,8	1998	7,3%	65,7	3136	8,3%	114,9	6006
SED	13,6%	58,3	4706	5,9%	49,9	2938	13,1%	184,1	14280
WM	15,6%	70,1	5314	5,8%	51,6	2134	7,3%	103,1	5170
WA	15,8%	70,9	5417	5,8%	51,2	2116	7,2%	102,2	5084
RN	5.63%	25.33	941	5.39%	25.1	848	5.22%	21.8	1.036

Fuente: Los autores

Cómo se observa en el cuadro anterior, los resultados de la red neuronal fueron más efectivos que los obtenidos con los modelos tradicionales en los tres años analizados, donde por ejemplo en el primer año el MAPE de la red fue de 5.63%, muy por debajo de 6.9% del SES que es el mejor de los tradicionales.

Para el segundo año, los valores se estrecharon más entre la red neuronal y los tradicionales, puesto que el MAPE para la red fue de 5.39%, mientras que en ambos Winters fue de 5.8% y en la SED fue de 5.9%; no obstante, al analizar el MAD y el MSE, la red neuronal si tuvo menores errores con un amplio margen respecto a los otros.

Para el tercer año pronosticado, el MAPE de la red neuronal (5.22%), fue menor que el generado por los demás métodos de pronósticos; los modelos de Winters Aditivo y Multiplicativo fueron los que más se le acercaron con 7.2% y 7.3% respectivamente.

Se observa además, como el error pronosticado de la red neuronal disminuyó ligeramente conforme se ampliaba el horizonte, pasando de 5.63% en el primer año a 5.39% en el segundo año y a 5.22% en el tercer año. Estas diferencias son de poca magnitud y lo relevante es que en los tres años evaluados la red neuronal generó mejores resultados que los modelos tradicionales analizados.

Es importante recordar que para la red desarrollada para pronosticar los casos nuevos de pacientes con cáncer hospitalizados al mes, se utilizaron algunas variables socioeconómicas, las cuales tenían un desfase de tres años, por lo tanto el pronóstico con este modelo va a tener ese techo superior de horizonte. No se puede ampliar el horizonte, ya que, cada tres años se deben actualizar dichas variables.

#### **4.6 Diseño de una herramienta en Excel para utilizar la red neuronal en proyecciones de pacientes con cáncer**

Como parte del desarrollo de la presente investigación se diseñó una herramienta automatizada para efectuar las proyecciones tanto para la red de casos nuevos como para la de estancia.

Esta herramienta, que fue desarrollada en Microsoft Excel 2007, incluye los procesos necesarios en cada una de las tres fases del modelo de la red neuronal, tanto para el aprendizaje, con base en los datos históricos; como la optimización, a través del algoritmo genético; y la operación, donde se generan los pronósticos, basándose en los datos de las variables de entrada.

En el caso de la herramienta para proyección de casos nuevos, está al utilizar los datos consolidados por mes, en el aprendizaje se incluyeron los datos de los últimos 14 años con que cuenta la Institución (de 1997 a 2010), no obstante, en el caso de la red de estancia, esta al considerar los casos individuales, se consolidó solamente con los últimos 1000 datos históricos, por cuanto las características de la herramienta hacen que no sea ágil utilizar un mayor volumen de datos.

En los Anexos 2 y 3 se presenta el manual de usuario de cada herramienta, donde se explica las consideraciones necesarias para su utilización, los diferentes procesos necesarios para efectuar el aprendizaje, la optimización y la generación de los pronósticos, de manera que se facilite su uso.

## CAPÍTULO V. APORTES TEÓRICOS Y METODOLÓGICOS DE LA INVESTIGACIÓN

Una vez desarrollada la investigación en forma completa, y analizados los diferentes factores que incidieron en el desarrollo de la red neuronal y en su optimización con algoritmos genéticos, se encontraron situaciones importantes desde el punto de vista teórico y metodológico que ayudaron a la conceptualización de los modelos, y que son importantes de resaltar en función de el uso de esta investigación, como referencia para la creación, por parte de otras personas, de modelos similares a los desarrollados.

Como primer aporte se tiene que para determinar las variables de entrada de un modelo de redes neuronales se debe analizar los efectos y relación de cada una de estas con respecto a la variable de salida. Se recomienda realizar análisis de correlaciones para determinar si existe dependencia significativa entre las variables que justifique su inclusión en el modelo.

Desde la perspectiva teórica, se considera la utilización del método de propagación del error hacia atrás (*back-propagation*) como la herramienta de aprendizaje de la red neuronal más utilizada para modelos de predicción, lo cual a través de esta investigación se pudo comprobar, ya que, al momento de la construcción de la red se analizaron otras opciones (por ejemplo aprendizaje puro con solo algoritmos genéticos), pero, se encontró que los mejores resultados se daban con este método.

Además se pudo comprobar que al utilizar un mayor número de datos para el aprendizaje, los resultados mejoran de manera importante. Lo anterior se determinó cuando se analizó la efectividad de la red neuronal de casos nuevos, ya que, cuando se analizó el horizonte de pronósticos se entrenó la red con diferente número de casos de entrada (144, 156,

168), y entre más casos se utilizaron para el entrenamiento menor fue el resultado del error de pronóstico.

En el caso de la red de estancia hospitalaria, debido a la limitación del programa computacional en que se desarrollo la aplicación, se tuvo que construir la red con 1000 casos, a pesar que se disponía de más de 70000. No obstante, con estos 1000 casos, se analizaron los errores construyendo el modelo con 250, 500, 750 y 1000 casos para el entrenamiento, y nuevamente entre más casos se utilizaron menor fue el valor del error de pronóstico.

De acuerdo con lo anterior no hay un tamaño de muestra específico para utilizar cuando se desarrollan modelos de redes neuronales, sino que depende de la aplicación específica, por ejemplo en el caso de pronósticos de demanda se recomienda utilizar como mínimo datos de 3 a 5 años relacionados con el fenómeno de estudio. En general como mínimo se recomienda utilizar datos que describan el comportamiento del fenómeno en al menos tres periodos de estudio, pero como se indicó, entre más datos se utilicen para el entrenamiento de la red neuronal mayor será el aprendizaje y mejor los resultados

Dentro del desarrollo de la red neuronal y la utilización del método de propagación del error hacia atrás (*back-propagation*), se tiene una tasa de aprendizaje de la red, esta tasa - aunque es solo un componente del diferencial- que se le suma a cada peso de la red al momento de hacer la propagación del error; al realizar las diferentes pruebas de los pesos se encontró que tiene un incidencia importante en la capacidad de la red de adaptarse al fenómeno, dado que, un pequeño cambio podía llevar a la red a crear errores tan grandes que hicieran el modelo no factible, es por ello que incluso se incluyó dentro del modelo de optimización de los algoritmos genéticos.

Otro aspecto a considerar en la construcción de un modelo de redes neuronales con muchas variables de entrada y varias neuronas ocultas (lo que desemboca en una cantidad importante de pesos dentro de la red) es la magnitud de los datos de entrada, ya que,

como se observó al crear el modelo de pronóstico de casos nuevos, si las variables se incluían con los valores con que estaban medidos, los pesos y la misma tasa de aprendizaje tenían que tener valores muy bajos para que los datos de salida fueran consistentes con los reales, datos del orden de  $10^{-12}$ , por lo que se optó por introducir las variables de entrada con un factor de multiplicación de  $10^{-3}$ , lo cual hizo que se mantuviera la proporcionalidad entre ellas y que los valores de los pesos y la tasa de aprendizaje sea más alta, y por lo tanto se aprecien mejor las diferencias.

Dentro de la creación de la red neuronal de pronóstico de casos nuevos de cáncer, se incluyó, aparte de las variables socioeconómicas que se analizaron anteriormente, algunas variables relacionadas con los datos reales, las cuales permitieron, de acuerdo con las pruebas realizadas, mejorar la efectividad de la red en cerca de un 15%, estas variables son los promedios de casos nuevos de los últimos tres años, dos años, un año, seis meses y tres meses, lo cual hizo que la red convergiera más fácilmente y mejorara el aprendizaje de las tendencias.

Dentro de la aplicación de los algoritmos genéticos, también se encontraron aspectos metodológicos importantes de resaltar, por ejemplo, se decidió incluir dos detalles con el fin de mejorar la convergencia de la red en un tiempo más corto, los cuales consistieron en hacer los cruces dando énfasis al mejor resultado, es decir cruzando el mejor con el segundo mejor en un primer cruce y posteriormente, cruzando el mejor con el tercer mejor, para el segundo cruce; adicionalmente se incluyó un cambio al momento de la mutación, dado que se conservaron sin mutar los dos mejores individuos.

En las pruebas realizadas se observó que, si bien es cierto, estos cambios producían que parezca más difícil explorar nuevas áreas de solución, se llegó a resultados aceptables y mejores que los otros métodos, en mucho menor tiempo que utilizando otras metodologías, expuestas en otras investigaciones, como solo conservar el mejor resultado o cruzando el mejor con el segundo mejor y el tercer mejor con el cuarto mejor.

Adicionalmente, y como medida de cierta contraposición de estas variaciones, se incluyó una tasa de mutación relativamente alta de 33%, lo cual hizo que el modelo pudiera explorar otras áreas de solución, a través de los seis individuos que sí participaron en la mutación.

Dentro de las especificaciones teóricas sobre los algoritmos genéticos, esta la aseveración de que a mayor número de iteraciones del proceso se pueden ir obteniendo mejores resultados, esta situación se pudo comprobar en el desarrollo de esta investigación, ya que, al aplicar más iteraciones del algoritmo se iban obteniendo mejores resultados, no obstante, también está claro que, como dice la teoría, entre más iteraciones se realicen más es el tiempo y los recursos invertidos, por lo tanto se debe tener un equilibrio entre el número de iteraciones y la mejora en los resultados, de manera que no se invierta mucho esfuerzo en mejoras en los resultados que no son significativas.

Otro aspecto importante dentro de la optimización con algoritmos genéticos que se logró identificar con el desarrollo de esta investigación, es que el error utilizado como función de optimización debe ser el error total de la red, incluyendo tanto la parte de la construcción como la de prueba el modelo, esto porque si solo se aplicara para un porcentaje de los datos, como los de prueba, la optimización tendía a llevar a la red a un aprendizaje dirigido solo a ese grupo de datos, y aunque se obtenían, para ese grupo de datos, errores de pronóstico muy pequeños, al momento de utilizar la red para casos diferentes, los resultados eran errores muy por encima, lo cual no ocurría al optimizar toda la red.

También es importante indicar que los datos de prueba de la red, al momento de optimizar con algoritmos genéticos, deben ser una parte importante de los datos totales, al menos superior al 15% de los mismos, ya que, según las pruebas realizadas si estos estaban por el orden del 10% o menos, los resultados de error al utilizar la red para seguir

estimando, no eran satisfactorios, caso contrario cuando se utilizó modelos con datos de prueba superiores al 15% o 20%.

Finalmente, con la presente investigación se determinó que utilizando los algoritmos genéticos es posible optimizar los pesos iniciales para las conexiones entre neuronas de la red, de modo que el error total de la red disminuyera. No obstante, la optimización tiene un efecto en la representación del fenómeno de estudio, ya que, tiende a disminuir la variabilidad u oscilación de los datos, y aunque en términos generales el error disminuía, no se lograba representar fielmente los “picos” que se presentan en algunos meses. Si la red no se optimizaba, es decir si únicamente se entrenaba con back-propagation, se lograba representar “los picos” y el comportamiento de los datos del fenómeno en estudio, pero, con una menor exactitud, lo que incrementaba el error de pronóstico.

## CONCLUSIONES

Mediante la presente investigación se desarrollaron y evaluaron dos modelos de redes neuronales, para proyecciones sobre pacientes con diagnóstico de cáncer hospitalizados en la CCSS. Se determinó que al optimizar ambas redes con algoritmos genéticos se generan mejores resultados, ya que, disminuye el error de pronóstico; en el caso de la red de casos nuevos este pasó de 7.53% a 5.07% después de optimizada, y en la red de estancia hospitalaria pasó de 47.37% a 40,79%.

La utilización de un modelo de redes neuronales, optimizadas con algoritmos genéticos, para proyectar la cantidad de casos nuevos de pacientes con cáncer hospitalizados en la CCSS, genera mejores resultados, en términos del error de pronóstico, que los obtenidos utilizando algunas de las técnicas tradicionales como suavización exponencial simple, suavización exponencial doble, winters aditivo winters multiplicativo. Mientras que con la red neuronal de casos nuevos se genera un error de pronóstico de 5.07%, el resultado más cercano con los métodos tradicionales indicados fue el de winters aditivo con 6.24% de error.

El modelo desarrollado presenta mejores resultados que los métodos tradicionales considerando un horizonte de pronósticos de uno, dos o tres años. Después de este período se recomienda hacer una actualización de datos de entrada de la red, así como un nuevo entrenamiento y optimización, ya que, los cambios en variables macroeconómicas y sociales incorporadas en el modelo se manifiestan con un rezago de tres años.

Las variables macroeconómicas y sociales del país pueden incidir en la cantidad de casos nuevos de pacientes con diagnóstico de cáncer hospitalizados en la CCSS. Aunque no se puede cuantificar la magnitud del aporte de cada variable al fenómeno de estudio, se determinó que al incluirlas en el modelo de redes neuronales con un rezago de tres años,

este genera mejores resultados de estimación de los casos nuevos, en términos del error de pronóstico.

Las variables relacionadas con las características y condiciones de los pacientes con cáncer hospitalizados en la CCSS, disponibles en la base de datos de egresos hospitalarios de dicha institución, son un insumo fundamental para desarrollar aplicaciones estadísticas, como por ejemplo las redes neuronales, que permitan una planificación y toma de decisiones oportuna respecto a los servicios que se le brindan a estos pacientes. Una de estas aplicaciones es el caso de la red neuronal para proyectar la estancia hospitalaria.

La construcción de un modelo de redes neuronales conlleva una comprensión profunda del fenómeno que se quiere representar, no solo desde el punto de vista matemático, sino desde el punto de vista lógico. La determinación del tipo de red, la cantidad de capas y del número de neuronas en la capa oculta, son elementos clave para la generación de resultados aceptables; aunque estos elementos algunas veces se determinan mediante prueba y error, hoy día se tienen a disposición programas informáticos estadísticos, que contienen módulos para modelar redes neuronales, que pueden utilizarse como guía para establecer los elementos en cuestión.

La aplicación automatizada del modelo de redes neuronales optimizadas con algoritmos genéticos, desarrollado como parte de este proyecto, además de facilitar el entrenamiento, optimización y análisis de resultados de las redes neuronales desarrolladas, permite que la CCSS la utilice para generar proyecciones de casos nuevos y de estancia hospitalaria, y que analice la conveniencia de implementarla formalmente.

La utilización de técnicas avanzadas de predicción y optimización, como lo son las redes neuronales y los algoritmos genéticos, permiten analizar otros factores que intervienen en los fenómenos en estudio, establecer relaciones que no parecen muy claras al entendimiento y razonabilidad matemática y lógica; sin embargo, a pesar de que se pueden crear estos modelos, la tecnología disponible para el manejo de los mismos, es

aún limitada, por lo que se debe restringir a modelos simples y con poco manejo de datos; lo cual con el desarrollo de las tecnologías y los programas computacionales, permitirá modificarse y avanzar en su uso y aprovechamiento en todos los sectores organizacionales.

## RECOMENDACIONES

A pesar de que existen programas informáticos que facilitan el modelado de las redes neuronales, estos son poco flexibles y no siempre permiten que el modelo sea una representación fidedigna de la realidad, por lo que, aunque puede utilizarse como guía para establecer algunos de los componentes de la red neuronal, se recomienda a quienes pretendan diseñar aplicaciones basadas en estas técnicas, que el modelo sea desarrollado manualmente en conjunto con las personas que conocen el fenómeno a estudiar, para así lograr una mejor representación del mismo.

Debido a las limitaciones observadas en los programas informáticos utilizados para la creación del modelo de redes neuronales, en cuanto al manejo de datos para el entrenamiento, se recomienda a la CCSS desarrollar una aplicación a la medida que permita entrenar la red de estancia con todos los casos disponibles, de modo que el proceso de aprendizaje sea más efectivo y se generen mejores resultados. Esto debe considerarse también para otras aplicaciones de redes neuronales que implique la utilización de la información sobre cada paciente individual.

La herramienta creada para el manejo de ambas redes neuronales, desarrollada sobre Microsoft Excel, es sencilla y de fácil utilización, sin embargo, la programación utilizada no es tan eficiente y la apariencia se puede mejorar. A pesar de esto, se recomienda que la CCSS la utilice como insumo para desarrollar una aplicación a la medida siguiendo la misma lógica de funcionamiento.

Las bases creadas con la presente investigación, en cuanto al manejo de ambas técnicas en dos temáticas propias de la CCSS, constituyen una metodología de desarrollo de nuevas aplicaciones, que permitan a la Institución aprovechar las ventajas de estas metodologías en otras aplicaciones de predicción, que ayuden a la planificación de los servicios en el corto y mediano plazo dentro de la Institución.

## REFERENCIAS BIBLIOGRÁFICAS

### Libros

- Adeli, H. & Hung, S. (1995). *Machine Learning: Neural Networks, Genetic Algorithms, and Fuzzy Systems*. New York, NY: John Wiley & Sons Inc.
- Angelini, J. (2000). *Pronósticos: Conceptos básicos*. San José, Costa Rica: Editorial Universidad de Costa Rica
- Freeman, J. y Skapura, D. (1993). *Redes neuronales: algoritmos, aplicaciones y técnicas de programación*. Argentina: Addison-Wesley.
- Grajales, T. (1996) *Conceptos Básicos para la Investigación Social*. Nuevo León, México: Publicaciones Universidad de Montemorelos.
- Hanke, J. (1992). *Pronósticos en los negocios* (5. ed.). México: Prentice Hall Hispanoamericana.
- Hilera, J. y Martínez, V. (2000). *Redes neuronales artificiales: fundamentos, modelos y aplicaciones*. México: Alfaomega.
- Jain Lakhmi C. y Martin, N.C. (1998). *Fusion of Neural Networks, Fuzzy Systems and Genetic Algorithms: Industrial Applications*. University of South Australia, Adelaide.
- Kurkova, V., Steele, N., Neruda, R. & Karny, M. (2001). *Artificial Neural Nets and Genetic Algorithms: Proceedings of the International Conference in Prague, Czech Republic*. Springer-Verlag New York, LLC
- Makridakis, S. (1992). *Pronósticos: Estrategia y planificación para el siglo XXI*. Madrid: Ediciones Díaz de Santos.
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1997). *Forecasting Methods and Applications*. New York: John Wiley & Sons, Inc.
- Masters, T. (1995). *Neural, Novel & Hybrid Algorithms for Time Series*. . New York, NY: John Wiley & Sons Inc.

- Reed, R. & Marks, R. (1999). *Neural Smothing: Supervised Learning in Feedforward Artificial Neural Networks*. Cambridge, Mass. : The MIT Press
- Rojas, R. (1996). *Neural Networks - A Systematic Introduction*. Springer-Verlag New York, LLC.
- Sapag Chain (2000). *Preparación y evaluación de proyectos*. Mc-Graw Hill.
- Solano, Y. (1997). *Improvement of convergence speed and generalization capability in artificial neural networks by normalizing the Back-propagation algorithm*. Chiba, Japan: s.n.e.
- Zhang, G., Patuwo, E & Hu, M. (1998). *Forecasting with artificial neural networks: The state of the art*. Graduate School of Management, Kent State University, Kent, Ohio. USA.
- Zhang, G. (2004). *Neural Networks in Business Forecasting*. Harrisburg, USA: Idea Group Publishing.

#### Tesis

- Camacho, S. (2001). *Propuesta de un modelo para la administración de ventas a corto plazo en Amanco de Costa Rica S.A.* Tesis para optar por Maestría en Administración de Negocios. Convenio UCR-FUNDEPOS.
- Calvo, A. (2004). *Un modelo de redes neuronales aplicado al Sistema Bancario Costarricense: período 1998-2002*. Tesis de Maestría, Facultad de Ciencias Económicas, Universidad de Costa Rica.
- Gallardo, E. (2009). *Modelo de redes neuronales artificiales aplicado a un problema de clasificación de estudiantes de secundaria según actitud hacia la aculturación*. Tesis de Maestría, Facultad de Ciencias Económicas, Universidad de Costa Rica.
- Orozco, C. y Medina, R. (2004). *Modelo de la evolución del conocimiento en individuos en un sistema autorregulado, basado en redes neuronales artificiales y*

*algoritmos genéticos*. Tesis para optar por el título de Ingeniero en Sistemas. Universidad EAFI. Medellín, Colombia. Disponible en la dirección <http://bdigital.eafit.edu.co/bdigital/PROYECTO/P006.32CDO747/capitulo1.pdf>

- Pacheco, J. Sancho, J y Zoch, A. (2006). *Rediseño de los procesos de administración de inventarios y almacenaje en la empresa Café El Rey*. Tesis para optar por el grado de Licenciatura en Ingeniería Industrial. Universidad de Costa Rica.
- Ramírez, J. (1999). *Ajuste de controladores difusos mediante algoritmos genéticos*. Tesis para optar por el grado de Magister Scientiae en Ingeniería Eléctrica. Universidad de Costa Rica.
- Ugalde, J. (1997). *Predicción de caudales en el manto acuífero herediano y consumos en la población de San Rafael de Heredia mediante redes neuronales*. Tesis de Licenciatura, Facultad de Ingeniería, Universidad de Costa Rica.
- Zúñiga, A. (2004). *Operación óptima de sistemas de potencia mediante algoritmos genéticos*. Tesis para optar por el grado de Magister Scientiae en Ingeniería Eléctrica con énfasis en Sistemas Digitales. Universidad de Costa Rica.

#### Artículos

- Autor corporativo Revista VIRTUALPRO (s.f.). *Pronósticos de Demanda*. Consultado el 6 de noviembre de 2010 en la dirección [http://www.revistavirtualpro.com/ediciones/Planeacion de la produccion Pronosticos de demanda-2010-09-01\\_5](http://www.revistavirtualpro.com/ediciones/Planeacion%20de%20la%20produccion%20Pronosticos%20de%20demanda-2010-09-01_5)
- Cevalloa, J. (2004) Aplicación de redes neuronales para optimizar problemas multirespuesta en mejora de la calidad. *Revista de la Facultad de Ingeniería Industrial*. Vol. 7 (2): pp. 31-34.
- Edelman, A. (s.f.). *Pronosticar la demanda: ¿Ejercicio de adivinación matemática o fundamento de la planeación operativa?*. Universidad de Montevideo. Consultado

el 6 de noviembre de 2010 desde la dirección [http://socrates.ieem.edu.uy/articulos/archivos/202\\_pronosticar\\_la\\_demanda.pdf](http://socrates.ieem.edu.uy/articulos/archivos/202_pronosticar_la_demanda.pdf).

- Gestal, M. (s.f.). *Introducción a los algoritmos genéticos*. Universidad de la Coruña. Consultado en la dirección electrónica <http://sabia.tic.udc.es/mgestal/cv/AAGGtutorial/TutorialAlgoritmosGeneticos.pdf>
- González, A. (2006). *Las mejores prácticas en pronósticos*. Corporate Resources Management, S.C. Disponible en la dirección <http://www.crmexico.com/articulos/Pronosticos.pdf>
- Herrera, F. Lozano, M. y Verdegay, J. (1994). *Algoritmos Genéticos: Fundamentos, Extensiones y Aplicaciones*. ETS de Ingeniería Informática. Universidad de Granada. España. Consultado en la dirección [ftp://decsai.ugr.es/pub/arai/tech\\_rep/lozano/esp/arbor.ps.Z](ftp://decsai.ugr.es/pub/arai/tech_rep/lozano/esp/arbor.ps.Z)
- Nojek, S., Britos, P., Rossi, B. y García Martínez, R. (2003). Pronóstico de Ventas: Comparación de Predicción basada en Redes Neuronales versus Método Estadístico. *Reportes Técnicos en Ingeniería del Software*. Vol 5 (1): pp. 1-12. Instituto Tecnológico de Buenos Aires-Argentina.
- Pita Fernández, S., Pértegas Díaz, S (2002). Investigación cuantitativa y cualitativa. Complejo Hospitalario-Universitario Juan Canalejo. España. Tomado de: [http://www.laalamedilla.org/Webpracticas/cd/documentos/cuanti\\_cuali2.pdf](http://www.laalamedilla.org/Webpracticas/cd/documentos/cuanti_cuali2.pdf)
- Sotolongo, G., Guzmán (2001), MV. Aplicaciones de las redes neuronales. El caso de la bibliometría. *Ciencias de la Información*. Vol 32(1): pp. 27-34. Consultado desde la dirección electrónica: <http://www.idict.cu/cinfo99/v32n1a2001/bibliometria.htm>
- Toro, O. Mejía, D. Salazar, H. Pronóstico de ventas usando redes neuronales. *Scientia Et Technica*. Vol 10(2): pp. 25-30. Universidad Tecnológica de Pereira-Colombia.

- Universidad del País Vasco. Facultad de Ciencias de la Computación. (s.f.) *Algoritmos Genéticos*. Consultado desde la dirección electrónica: <http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/temageneticos.pdf>
- Valenzuela, M (2008). *Logística Inteligente*. Centro de Computación Inteligente y Robótica Tecnológico de Monterrey, Campus Monterrey. Disponible en [http://strauss.ing.puc.cl/laccir/Portals/0/research\\_teams/itesm/optimization\\_and\\_logistics.pdf](http://strauss.ing.puc.cl/laccir/Portals/0/research_teams/itesm/optimization_and_logistics.pdf)

#### Otros

- Ramírez, P. (2004). *Pronósticos de demanda*. San José, Costa Rica: Universidad de Costa Rica. Material del curso Diseño de Planta.

# ANEXOS

## Anexo 1: Resultados de los pronósticos utilizando métodos tradicionales mediante el programa Minitab

### Suavización exponencial simple para Casos nuevos pacientes cáncer

Datos Casos nuevos pacientes cáncer  
Longitud 156

Constante de suavización

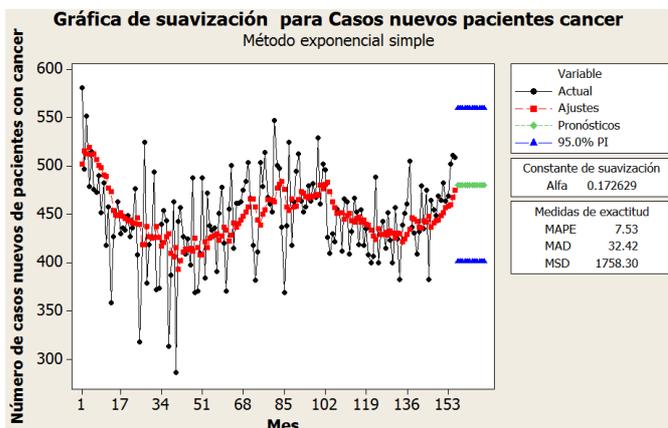
Alfa 0.172629

Medidas de exactitud

MAPE 7.53  
MAD 32.42  
MSD 1758.30

Pronósticos

Período	Pronóstico	Inferior	Superior
157	480.790	401.362	560.219
158	480.790	401.362	560.219
159	480.790	401.362	560.219
160	480.790	401.362	560.219
161	480.790	401.362	560.219
162	480.790	401.362	560.219
163	480.790	401.362	560.219
164	480.790	401.362	560.219
165	480.790	401.362	560.219
166	480.790	401.362	560.219
167	480.790	401.362	560.219
168	480.790	401.362	560.219



## Suavización exponencial doble para Casos nuevos pacientes cáncer

Datos Casos nuevos pacientes cáncer  
 Longitud 156

Constantes de suavización

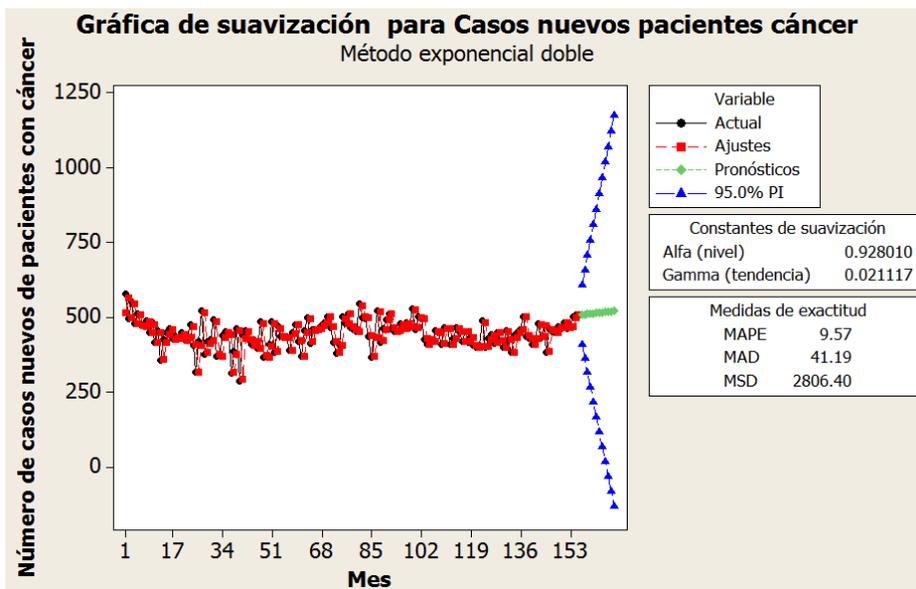
Alfa (nivel) 0.928010  
 Gamma (tendencia) 0.021117

Medidas de exactitud

MAPE 9.57  
 MAD 41.19  
 MSD 2806.40

Pronósticos

Periodo	Pronóstico	Inferior	Superior
157	510.289	409.371	611.21
158	511.404	364.418	658.39
159	512.519	316.753	708.28
160	513.634	267.985	759.28
161	514.749	218.673	810.82
162	515.864	169.053	862.67
163	516.978	119.243	914.71
164	518.093	69.307	966.88
165	519.208	19.285	1019.13
166	520.323	-30.800	1071.45
167	521.438	-80.931	1123.81
168	522.553	-131.099	1176.21



## Método de Winters para Casos nuevos pacientes cáncer

Método aditivo

Datos Casos nuevos pacientes cáncer  
Longitud 156

Constantes de suavización

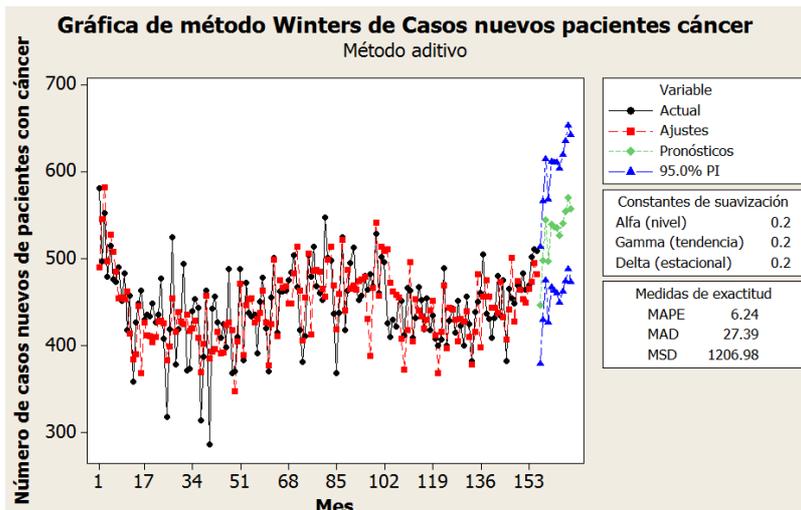
Alfa (nivel) 0.2  
Gamma (tendencia) 0.2  
Delta (estacional) 0.2

Medidas de exactitud

MAPE 6.24  
MAD 27.39  
MSD 1206.98

Pronósticos

Periodo	Pronóstico	Inferior	Superior
157	446.590	379.483	513.698
158	498.255	430.096	566.413
159	544.824	475.493	614.155
160	497.533	426.915	568.151
161	539.678	467.664	611.691
162	536.632	463.120	610.145
163	535.158	460.050	610.265
164	526.631	449.839	603.423
165	540.798	462.236	619.361
166	554.637	474.227	635.048
167	570.370	488.036	652.703
168	557.410	473.086	641.735



## Método de Winters para Casos nuevos pacientes cáncer

Método multiplicativo

Datos Casos nuevos pacientes cáncer  
Longitud 156

Constantes de suavización

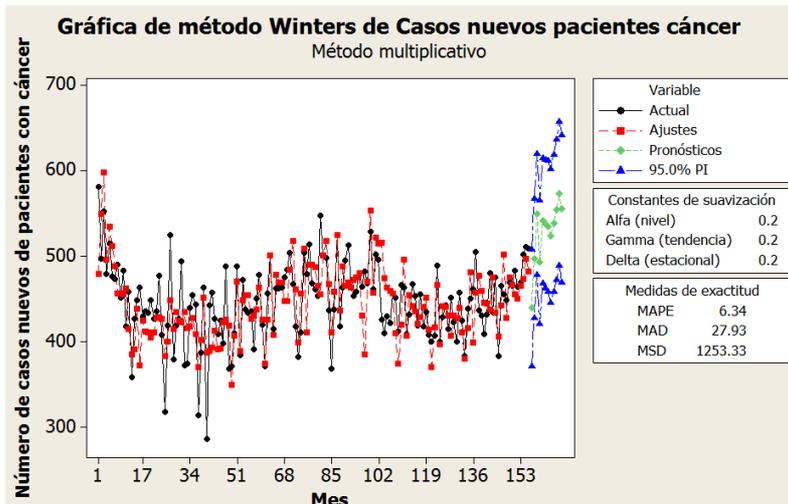
Alfa (nivel) 0.2  
Gamma (tendencia) 0.2  
Delta (estacional) 0.2

Medidas de exactitud

MAPE 6.34  
MAD 27.93  
MSD 1253.33

Pronósticos

Periodo	Pronóstico	Inferior	Superior
157	439.580	371.143	508.018
158	497.210	427.700	566.719
159	549.210	478.505	619.915
160	493.303	421.285	565.321
161	541.304	467.863	614.746
162	537.684	462.714	612.653
163	534.952	458.356	611.548
164	523.750	445.436	602.065
165	538.844	458.725	618.964
166	553.995	471.990	635.999
167	573.222	489.257	657.187
168	555.569	469.574	641.565



**Anexo 2: Manual de Usuario de la Herramienta de pronósticos de casos nuevos de  
cáncer atendidos en los centros de la CCSS**

# **Manual de Apoyo**

## **Herramienta de pronóstico de casos nuevos de cáncer hospitalizados en la CCSS**

### **Utilizando Redes Neuronales optimizadas con Algoritmos Genéticos**

Andrey Alonso Mora Bermúdez

Juan Antonio Sancho Chacón

**2011**

## **Introducción**

En este manual se presentan las principales características de la Herramienta de pronóstico de casos nuevos de cáncer hospitalizados en la CCSS; dicha herramienta está basada en un modelo de pronóstico desarrollado con Redes Neuronales, optimizada con Algoritmos Genéticos; de forma que el usuario final pueda conocer su funcionamiento, capacidad y limitaciones generales.

El instrumento consiste en un libro de Excel que, a través de macros escritas en VBA (Visual Basic Applications), procesa datos provenientes del mismo libro (archivos fuente) de una manera sencilla y amigable para el usuario.

El instrumento se encuentra dividido en cuatro módulos básicos, cada uno de los cuales realiza una determinada tarea de procesamiento y archivo de datos para obtener así las funcionalidades necesarias para obtener la proyección de casos nuevos de cáncer hospitalizados en la CCSS. Cada uno de los módulos desarrollados en la Herramienta será explicado más adelante en éste Manual.

## **Requisitos mínimos**

La ejecución apropiada del instrumento presenta los siguientes requisitos mínimos:

- Hardware: Procesador 2 GHz, 2Gb de memoria RAM.
- Sistema operativo: Microsoft Windows Vista (cualquier versión).
- Software: Microsoft Excel 2007 para Windows o posterior, instalación completa (se debe habilitar la ejecución de macros).
- Resolución de pantalla: 1024 x 768 o superior.

## Acerca de la interfaz de usuario

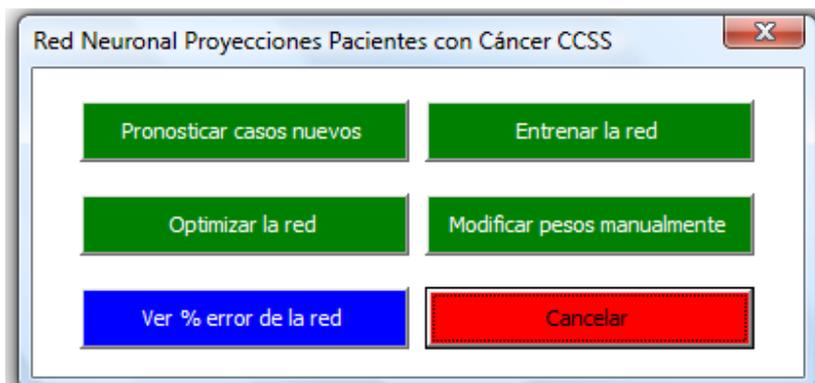
Se ha procurado emplear una interfaz amigable a lo largo del instrumento para facilitar su utilización, la navegación entre módulos y tareas, y la apreciación visual de la información, que en suma propicie la toma correcta de decisiones. Esta interfaz se encuentra compuesta por una página de inicio, la cual posee el nombre de la Herramienta y los autores. La interfaz se presenta en la Imagen 1.

### Imagen 1. Página inicial de la herramienta



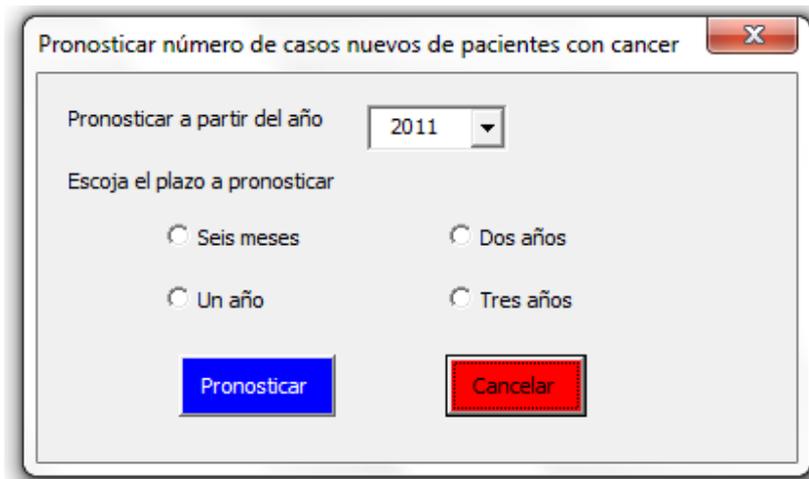
Posteriormente haciendo clic en el ícono Inicio, se puede acceder al menú de módulos, desde donde se puede acceder a cada uno de los 4 módulos, o devolverse a la interfaz de inicio. Este menú de módulos se presenta en la imagen 2.

### Imagen 2. Menú de módulos



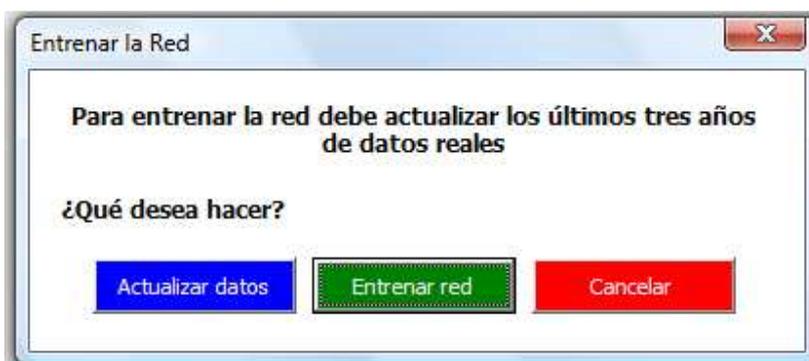
El módulo de pronosticar casos nuevos, presenta una interfaz donde el usuario puede elegir diferentes períodos de pronóstico, contabilizados a partir de un año en específico. La interfaz se presenta en la imagen 3.

**Imagen 3. Interfaz de pronóstico de casos nuevos**



El siguiente módulo es el de entrenamiento de la red neuronal, al acceder a este módulo se presenta una pantalla con tres opciones, las de actualizar datos para poder entrenar, la de ejecutar el entrenamiento y la de regresar, esta pantalla se observa en la imagen 4.

**Imagen 4. Interfaz de entrenamiento de la red neuronal**



El siguiente módulo es el de optimización de la red neuronal, accediendo a este módulo se presenta una página de Excel donde se puede observar el valor inicial de los pesos de

la red neuronal, así como los resultados en cuanto al error de construcción, el de prueba y el error total de la red; así mismo tiene la opción de ir haciendo cada paso del proceso de optimización, así como iniciar la optimización con 45 iteraciones o agregar 45 iteraciones al proceso. Esta pantalla se presenta en la imagen 5.

**Imagen 5. Pantalla de optimización de la red**

**OPTIMIZAR PESOS W**

Error prueba	Error const	Error total %	F. Aprendizaje	V1							V2								
				O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	
5.45%	7.13%	6.77%	7.7638E-06	75.0711	13.9682	-39.78	51.076	17.918	-44.573	7.4188	-21.269	44.154	7.2404	-24.407	-19.022	-36.275	-39.015	-12.885	
		6.87%	7.7638E-06	75.0711	13.9682	-39.78	51.076	17.918	-44.573	7.4188	-21.269	44.154	7.2404	-24.407	-19.022	-36.275	-39.015	-12.885	
		7.00%	1.6277E-05	71.6249	12.1216	-39.129	41.207	14.539	-40.499	6.9755	-22.581	44.716	7.1119	-24.408	-19.137	-36.074	-38.872	-12.087	
<b>Aleatorio Inicial</b>				5.3699E-05	-8.13552	-6.23759	8.36788	-7.6981	16.514	14.71	41.013	-19.248	-43.083	28.38	-30.677	23.213	-3.1043	2.4065	11.999

PRIMERA GENERACION

El cuarto y último modo corresponde a una opción que tiene el instrumento para intervenir en el accionar de la red neuronal, el mismo consiste en la modificación manual de los pesos de la red, así como del factor de aprendizaje. Para esto se accede a una pantalla donde se pueden incluir los valores y luego agregarlo al modelo. La pantalla de acceso se muestra en la imagen 6.

**Imagen 6. Pantalla de ejecución de la modificación manual de los pesos**

**Modificar pesos de la red manualmente**

**Instrucciones**  
 Introduzca los nuevos valores de frecuencia de aprendizaje, pesos de las neuronas de entrada y de las neuronas ocultas. Luego presione el botón azul de la parte inferior de la pantalla.

Frecuencia de aprendizaje

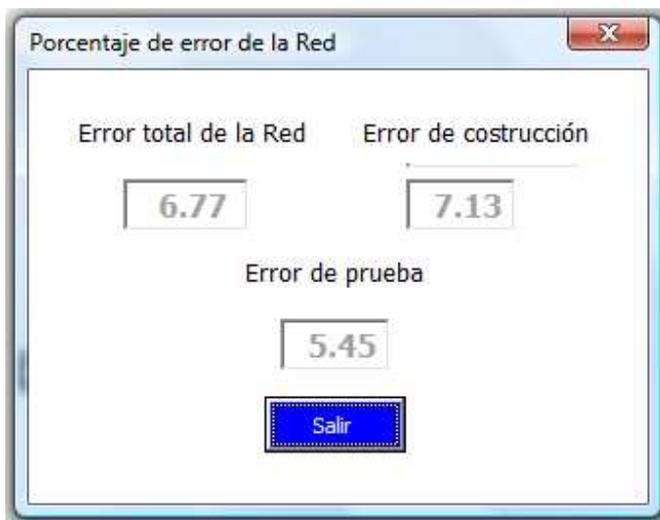
Sexo							Promedio 3 años							Promedio 2 años							Promedio 1 año							Promedio 6 meses							Promedio 3 meses												
O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7						
V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26	V27	V28	V29	V30	V31	V32	V33	V34	V35	V36	V37	V38	V39	V40	V41	V42	V43						

Tasa general de mortalidad							Indice de Desarrollo Humano							Inversión Social en Salud (% del PIB)							Cobertura Regimen de Enfermedad y Maternidad							Tasa de desempleo							Nivel de pobreza						
O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7
V44	V45	V46	V47	V48	V49	V50	V51	V52	V53	V54	V55	V56	V57	V58	V59	V60	V61	V62	V63	V64	V65	V66	V67	V68	V69	V70	V71	V72	V73	V74	V75	V76	V77	V78	V79	V80	V81	V82	V83	V84	V85

Neuronas ocultas						
O1	O2	O3	O4	O5	O6	O7
V86	V87	V88	V89	V90	V91	V92

Finalmente, en el menú de módulos, se puede acceder a una consulta directa a los errores de la red, que son los que dictan la eficacia de la misma, este acceso permite conocer en cualquier momento los resultados del último aprendizaje de la red, con el grupo de pesos que están actualmente vigentes, la pantalla que se activa se presenta en la siguiente imagen.

**Imagen 7. Resultados de error de la red**



En esta se puede consultar el error total de la red, el error de construcción y el error de prueba, con lo que se puede validar la eficacia de la misma. Los errores presentados están en términos de porcentajes.

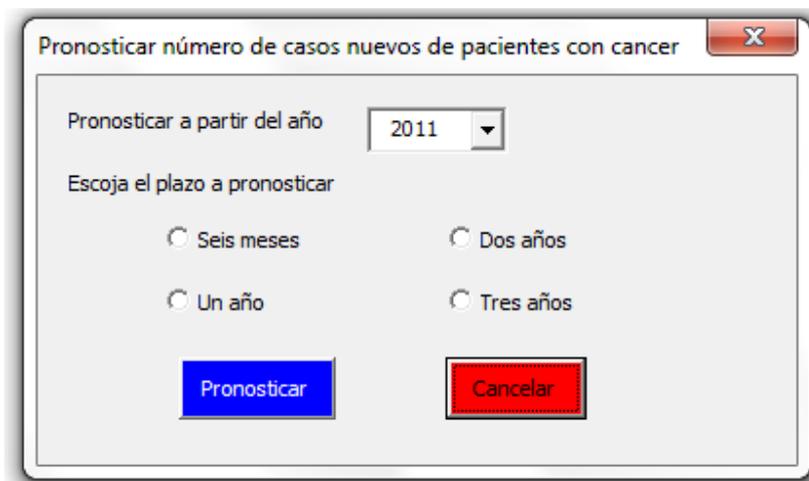
A continuación se presenta una descripción del manejo y la funcionalidad de cada uno de los módulos que componen la herramienta.

## Módulos

### 1. Pronosticar casos nuevos

El módulo de Pronosticar casos nuevos es el que permite hacer proyecciones de los casos nuevos de pacientes con cáncer hospitalizados en la CCSS que se estiman se presentarán en los próximos años. La interfaz de este módulo se presenta en la imagen 8.

**Imagen 8. Interfaz del módulo de pronosticar casos nuevos**



Esta interfaz presenta una serie de componentes que se describen a continuación:

- 1. Escogencia del año a partir del cual generar el pronóstico:** Este permite escoger el año a partir del cual se pueden ver las estimaciones, partiendo del año 2011 y hasta el 2018.
- 2. Período a pronosticar:** Los pronósticos pueden ser para períodos de 6 meses, un año, dos años y tres años, que es el máximo del período a pronosticar, debido a que se utiliza un desfase en las variables macroeconómicas, que sirven como entrada de la red, de 3 años. Es por ello que la herramienta tiene una validez inicial de generar pronósticos hasta el año 2020.
- 3. Botón de Pronosticar:** El cual permite ver los resultados del pronóstico para el período escogido.
- 4. Botón de Cancelar:** Este devuelve la herramienta a la pantalla del menú de módulos.

Para observar los resultados del pronóstico se accede a una pantalla donde se presentan el período o períodos anuales escogido, con un total de casos por año o período de seis meses, tal como se muestra en la imagen 9.

**Imagen 9. Resultado del pronóstico**

**Pronóstico**

[Regresar al Menú](#)

Mes	Número de casos nuevos	Mes	Número de casos nuevos	Mes	Número de casos nuevos
ene-11	436	ene-12	439	ene-13	444
feb-11	449	feb-12	445	feb-13	446
mar-11	456	mar-12	449	mar-13	448
abr-11	454	abr-12	451	abr-13	453
may-11	460	may-12	456	may-13	457
jun-11	463	jun-12	460	jun-13	462
jul-11	472	jul-12	464	jul-13	466
ago-11	471	ago-12	465	ago-13	468
sep-11	469	sep-12	465	sep-13	470
oct-11	471	oct-12	467	oct-13	471
nov-11	470	nov-12	469	nov-13	474
dic-11	470	dic-12	472	dic-13	476
<b>Total</b>	<b>5541</b>	<b>Total</b>	<b>5502</b>	<b>Total</b>	<b>5535</b>

Esta pantalla tiene un botón de Regresar al Menú, donde se puede regresar al menú de módulos, para continuar viendo otras estimaciones o accediendo a otros módulos.

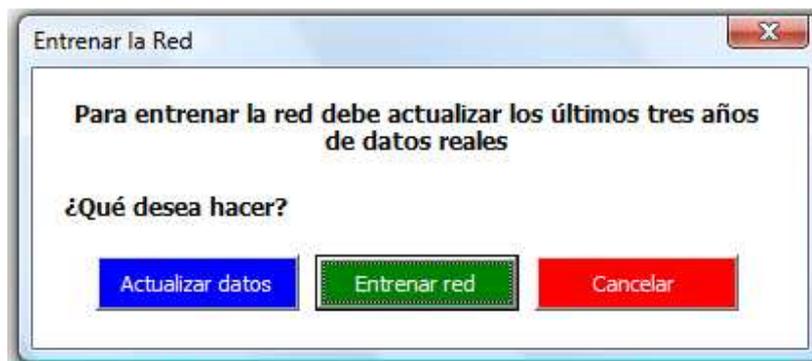
## **2. Entrenar la Red**

Este modulo permite ejecutar el proceso de entrenamiento de la red neuronal, este proceso consiste en determinar un nuevo conjunto de pesos para las relaciones entre variables, que mejore lo resultados de los pronósticos, se utiliza generalmente cuando, al pasar el tiempo, hay un cambio importante en los datos reales, que puedan afectar los resultados de la red.

El proceso está definido en la herramienta para que se ejecute una sola vez cada 3 años, una vez que se han actualizado los datos reales de las variables de entrada de esos tres años y es un proceso irreversible.

La pantalla de inicio del módulo se presenta en la imagen 10.

#### Imagen 10. Interfaz del módulo de Entrenar la red



El módulo presenta una interfaz donde con 4 diferentes elementos:

- 1. Una leyenda:** En esta se menciona la necesidad de actualizar los tres años de datos reales, necesarios para la optimización. La red está inicialmente entrenada con los datos reales de los casos nuevos hasta diciembre de 2010, utilizando las variables macroeconómicas de entrada correspondientes, hasta diciembre de 2007. Adicionalmente la red ya contiene los valores reales de las variables macroeconómicas hasta 2010, con lo que se puede tramitar el siguiente entrenamiento, una vez que se actualicen los datos reales de pacientes nuevos con cáncer hospitalizados en la CCSS hasta diciembre de 2013.
- 2. Botón de Actualizar datos:** Este permite al usuario acceder a la pantalla donde se presentan los datos reales, tanto de los casos nuevos de cáncer como de las variables macroeconómicas, en donde se puede ingresar los datos reales de las mismas. Este proceso se puede hacer en cualquier momento y con base en los cambios plasmados, se van ir ajustando los datos de los pronósticos; no obstante, para efectuar el entrenamiento formal si se debe completar los datos

del trienio correspondiente. La pantalla de esta página se presenta en la imagen 11.

**Imagen 11. Pantalla de datos reales**

Datos reales														
<a href="#">Regresar al Menú</a>														
	Casos Reales	Mes	Promedio 3 años	Promedio 2 año	Promedio 1 año	Promedio 6 meses	Promedio 3 meses		Tasa general de mortalidad	Indice de Desarrollo Humano	Social en Salud (% del PIB)	Cobertura Regimen de Enfermedad y Maternidad	Tasa de desempleo	Nivel de pobreza
ene-97	581	1	581	581	581	581	581	ene-94	4	0.85	4.8	86.2	4.2	20.0
feb-97	497	2	581	581	581	581	581	feb-94	4	0.85	4.8	86.2	4.2	20.0
mar-97	552	3	539	539	539	539	539	mar-94	4	0.85	4.8	86.2	4.2	20.0
abr-97	479	4	543	543	543	543	543	abr-94	4	0.85	4.8	86.2	4.2	20.0
may-97	515	5	527	527	527	527	509	may-94	4	0.85	4.8	86.2	4.2	20.0
jun-97	476	6	525	525	525	525	515	jun-94	4	0.85	4.8	86.2	4.2	20.0
jul-97	473	7	517	517	517	517	490	jul-94	4	0.85	4.8	86.2	4.2	20.0
ago-97	490	8	510	510	510	499	488	ago-94	4	0.85	4.8	86.2	4.2	20.0
sep-97	452	9	508	508	508	498	480	sep-94	4	0.85	4.8	86.2	4.2	20.0
oct-97	483	10	502	502	502	481	472	oct-94	4	0.85	4.8	86.2	4.2	20.0
nov-97	418	11	500	500	500	482	475	nov-94	4	0.85	4.8	86.2	4.2	20.0
dic-97	458	12	492	492	492	465	451	dic-94	4	0.85	4.8	86.2	4.2	20.0
ene-98	359	1	490	490	490	462	453	ene-95	4.2	0.85	4.6	86.4	5.2	20.4
feb-98	427	2	479	479	471	443	412	feb-95	4.2	0.85	4.6	86.4	5.2	20.4
mar-98	449	3	476	476	465	433	415	mar-95	4.2	0.85	4.6	86.4	5.2	20.4
abr-98	463	4	474	474	457	432	412	abr-95	4.2	0.85	4.6	86.4	5.2	20.4
may-98	430	5	473	473	455	429	446	may-95	4.2	0.85	4.6	86.4	5.2	20.4

En esta pantalla los datos que se pueden modificar por el usuario son los que están marcados en verde, correspondientes a los casos reales de casos nuevos de pacientes con cáncer hospitalizados en la CCSS y las variables macroeconómicas, las cuales se encuentran con una escala diferente de tiempo, para que queden alineadas tal como se utilizan en el modelo.

Adicionalmente la pantalla tiene un botón de regresar al menú, el cual permite volver al menú de módulos.

Es importante aclarar que esta actualización de datos no se debe hacer solamente cuando se va a optimizar la red; sino que cada vez que se tengan los datos reales de un período específico se puede actualizar, de manera que la red pueda ajustar el pronóstico, al menos por el efecto en las variables de entrada.

- 3. Botón de Entrenar la Red:** Este botón permite ejecuta el entrenamiento de la red para el nuevo período, una vez que se ejecuta, se le consulta al usuario si se está seguro de querer efectuar el proceso, esta es una validación para evitar errores de iniciar el proceso sin querer. Una vez que este acepte la ejecución del

proceso el mismo se efectúa y envía un mensaje de que el proceso fue exitoso, posteriormente se presenta la pantalla de resultados de error de la red, con lo cual el usuario puede ver los resultados obtenidos en el aprendizaje.

- Botón de Cancelar:** Este devuelve la herramienta a la pantalla del menú de módulos.

### 3. Optimizar la red

El módulo de optimizar la red permite acceder a una pantalla donde se puede hacer el proceso de optimización.

Esta pantalla permite efectuar la optimización de la red de tres formas diferentes, un proceso manual de inicio, compuesto por 18 pasos que conforman 3 iteraciones; un proceso automático que inicia el proceso de optimización con 45 iteraciones y un proceso que agrega 45 iteraciones, pero, conservando los datos que se tengan hasta el momento. La pantalla se muestra en la imagen 12, posteriormente se explican cada una de las partes.

**Imagen 12. Pantalla de optimización de la red**

**OPTIMIZAR PESOS W**

Limpiar pantalla
Regresar al Menú
Iniciar 45 iteraciones
Aplicar 45 iteraciones adicionales

Error prueba	Error construcción	Error total %	F. Aprendizaje	V1							V2							
				O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	
5.45%	7.13%	6.77%	7.7638E-06	75.0711	13.9682	-39.78	51.076	17.918	-44.573	7.4188	-21.269	44.154	7.2404	-24.407	-19.022	-36.275	-39.015	
			6.87%	7.7638E-06	75.0711	13.9682	-39.78	51.076	17.918	-44.573	7.4188	-21.269	44.154	7.2404	-24.407	-19.022	-36.275	-39.015
			7.00%	1.6277E-05	71.6249	12.1216	-39.129	41.207	14.539	-40.499	6.9755	-22.581	44.716	7.1119	-24.408	-19.137	-36.074	-38.872
			Aleatorio Inicial	5.8446E-06	21.0113	24.0842	49.1147	-17.189	-42.264	23.709	8.2245	12.585	-32.17	-43.809	44.158	7.6856	9.3876	-39.625

**PRIMERA GENERACION**

Crear población inicial

Cromosoma	Error	FA	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14
1	37.84%	3.9021E-06	16.2179	26.3558	-39.851	-34.819	-22.999	17.519	27.709	11.596	-26.661	9.1898	33.121	-25.385	-43.635	-21.078
2	11.41%	8.1466E-05	-18.4341	41.8123	-8.9586	3.9862	13.479	22.786	26.611	7.8156	1.1575	-49.166	-34.472	10.328	34.199	43.28
3	14.08%	4.2743E-05	28.3967	40.2865	-43.977	-1.2072	30.29	23.683	33.484	-31.614	33.899	41.579	-32.531	47.602	-29.209	-9.8721
4	14.10%	9.3229E-05	36.1982	-25.2939	-36.063	-37.224	-39.774	-48.826	43.398	-49.25	-49.285	-30.095	-5.2756	8.0255	-48.473	-20.664
5	20.66%	6.6391E-05	41.6582	25.9795	-16.207	-12.427	9.6238	20.197	45.745	-45.103	43.437	48.855	-3.6916	44.644	16.671	39.54
6	14.46%	8.2E-05	-12.5256	7.70991	-49.839	23.647	-3.8787	25.503	-5.2784	-25.575	48.268	21.319	-4.965	10.689	-40.042	-49.5
7	14.13%	9.9005E-05	43.3392	31.3175	10.9199	19.319	-17.356	-38.587	48.094	-30.832	-39.207	37.341	-21.894	-37.065	-41.155	-8.3506
8	14.50%	4.6955E-05	-36.8867	44.4523	-10.933	-30.961	11.267	-34.707	-15.23	38.559	45.111	-24.866	25.39	-24.094	-0.4332	-31.611

Ordenar población inicial

1. **Botón Limpiar pantalla:** Este botón permite limpiar todos los datos de las diferentes tablas que conforman el proceso de optimización.
2. **Botón Regresar al Menú:** Este devuelve la herramienta a la pantalla del menú de módulos.
3. **Botón de Iniciar 45 iteraciones:** Este botón permite hacer un proceso automático de aplicación de 45 iteraciones (rutinas), las cuales iniciar desde cero, es decir que se construyen nuevos individuos (conjuntos de datos) y se van realizando los diferentes procesos para la optimización.
4. **Botón de Aplicar 45 iteraciones adicionales:** Este botón permite agregar 45 iteraciones de optimización adicionales, que permiten mejorar los resultados que ya se tienen, es decir arranca con un grupo de individuos que ya fueron utilizados anteriormente y sobre esos es que se van mejorando. Este proceso se utiliza normalmente después del de iniciar 45 iteraciones, o después de correr el proceso manual de tres iteraciones de optimización.
5. **Proceso manual de optimización:** Este proceso consiste en aplicar tres rutinas de optimización, pero, haciéndolo paso a paso, mediante 17 botones de color azul que se presentan en la pantalla, los cuales deben ser aplicados en el orden respectivo e inician desde cero, generando una nueva población de análisis (grupo de soluciones) y mejorando a partir de ella.
6. **Información de la red:** La pantalla también tiene un espacio donde se muestra los resultados del error de pronóstico de construcción, prueba y total de la red, según la población que se esté evaluando, esto permite observar mejor el comportamiento de la optimización, para asegurarse que los resultados mejoran el accionar de la red, especialmente en cuanto al error total y el de prueba.

#### **4. Modificar pesos manualmente**

Este módulo consiste en una pantalla que permite modificar los pesos de las diferentes conexiones de forma manual, así como la frecuencia de aprendizaje, de manera que un

usuario experto, puede hacer ajustes en los pesos, buscando mejorar los resultados. La pantalla se muestra en la imagen 13.

Imagen 13. Pantalla de modificación manual de pesos

### Modificar pesos de la red manualmente

**Instrucciones**      Introduzca los nuevos valores de frecuencia de aprendizaje, pesos de las neuronas de entrada y de las neuronas ocultas. Luego presione el botón azul de la parte inferior de la pantalla.

Frecuencia de aprendizaje

Sexo							Promedio 3 años							Promedio 2 años							Promedio 1 año							Promedio 6 meses							Promedio 3 meses						
O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7
V/2	V/3	V/4	V/5	V/6	V/7	V/8	V/9	V/10	V/11	V/12	V/13	V/14	V/15	V/16	V/17	V/18	V/19	V/20	V/21	V/22	V/23	V/24	V/25	V/26	V/27	V/28	V/29	V/30	V/31	V/32	V/33	V/34	V/35	V/36	V/37	V/38	V/39	V/40	V/41	V/42	V/43

Tasa general de mortalidad							Índice de Desarrollo Humano							Inversión Social en Salud (% del PIB)							Cobertura Regimen de Enfermedad y Maternidad							Tasa de desempleo							Nivel de pobreza						
O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7
V/44	V/45	V/46	V/47	V/48	V/49	V/50	V/51	V/52	V/53	V/54	V/55	V/56	V/57	V/58	V/59	V/60	V/61	V/62	V/63	V/64	V/65	V/66	V/67	V/68	V/69	V/70	V/71	V/72	V/73	V/74	V/75	V/76	V/77	V/78	V/79	V/80	V/81	V/82	V/83	V/84	

Neuronas ocultas						
O1	O2	O3	O4	O5	O6	O7
V/88	V/87	V/88	V/89	V/90	V/91	V/92

Modificar Factor de Aprendizaje y Pesos

Regresar al Menú

Esta pantalla está compuesta por las siguientes partes:

1. **Área de introducción de pesos:** Está compuesta por 92 espacios para modificar, 84 de conexiones de las variables de entrada con las variables ocultas y 7 de las ocultas con la de salida, además de un espacio para modificar la frecuencia de aprendizaje.
2. **Botón Modificar Factor de Aprendizaje y Pesos:** Este botón permite realizar la sustitución de los datos compilados en los espacios, sobre la red neuronal. Cabe mencionar que este proceso modifica todos los datos, lo cual implica que si un dato está en blanco lo sustituirá por cero en el modelo, por lo tanto la modificación de pesos se debe realizar actualizando todos los pesos.
3. **Botón de Regresar al Menú:** Este devuelve la herramienta a la pantalla del menú de módulos.

**Anexo 3: Manual de Usuario de la Herramienta de pronósticos de estancia hospitalaria  
de los pacientes con cáncer atendidos en la CCSS**

# **Manual de Apoyo**

## **Herramienta para pronóstico de estancia hospitalaria de pacientes con cáncer**

### **Utilizando Redes Neuronales optimizadas con Algoritmos Genéticos**

Andrey Alonso Mora Bermúdez

Juan Antonio Sancho Chacón

**2011**

## **Introducción**

En este manual se presentan las principales características de la Herramienta de pronóstico de estancia hospitalaria para pacientes con cáncer hospitalizados en la CCSS; dicha herramienta está basada en un modelo de pronóstico desarrollado con Redes Neuronales, optimizada con Algoritmos Genéticos; de forma que el usuario final pueda conocer su funcionamiento, capacidad y limitaciones generales.

El instrumento consiste en un libro de Excel que, a través de macros escritas en VBA (Visual Basic Applications), procesa datos provenientes del mismo libro (archivos fuente) de una manera sencilla y amigable para el usuario.

El instrumento se encuentra dividido en cuatro módulos básicos, cada uno de los cuales realiza una determinada tarea de procesamiento y archivo de datos para obtener así las funcionalidades necesarias para obtener la proyección de estancia hospitalaria de los pacientes con cáncer. Cada uno de los módulos desarrollados en la Herramienta será explicado más adelante en éste Manual.

## **Requisitos mínimos**

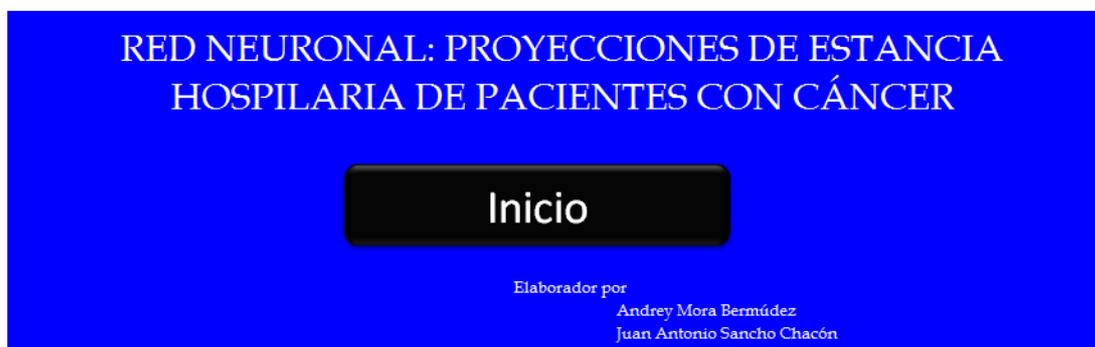
La ejecución apropiada del instrumento presenta los siguientes requisitos mínimos:

- Hardware: Procesador 2 GHz, 2Gb de memoria RAM.
- Sistema operativo: Microsoft Windows Vista (cualquier versión).
- Software: Microsoft Excel 2007 para Windows o posterior, instalación completa (se debe habilitar la ejecución de macros).
- Resolución de pantalla: 1024 x 768 o superior.

## Acerca de la interfaz de usuario

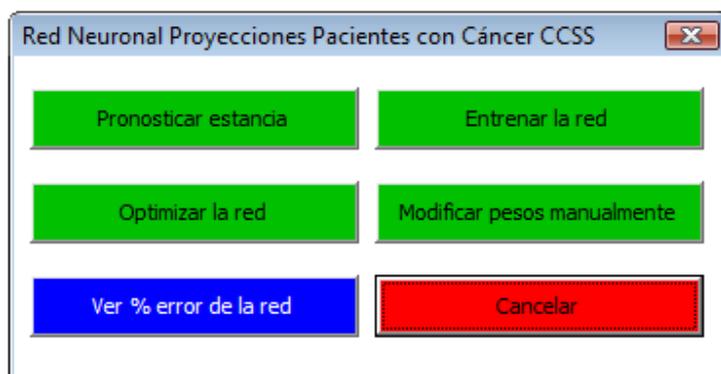
Se ha procurado emplear una interfaz amigable a lo largo del instrumento para facilitar su utilización, la navegación entre módulos y tareas, y la apreciación visual de la información, que en suma propicie la toma correcta de decisiones. Esta interfaz se encuentra compuesta por una página de inicio, la cual posee el nombre de la Herramienta y los autores. La interfaz se presenta en la Imagen 1.

**Imagen 1. Página inicial de la herramienta**



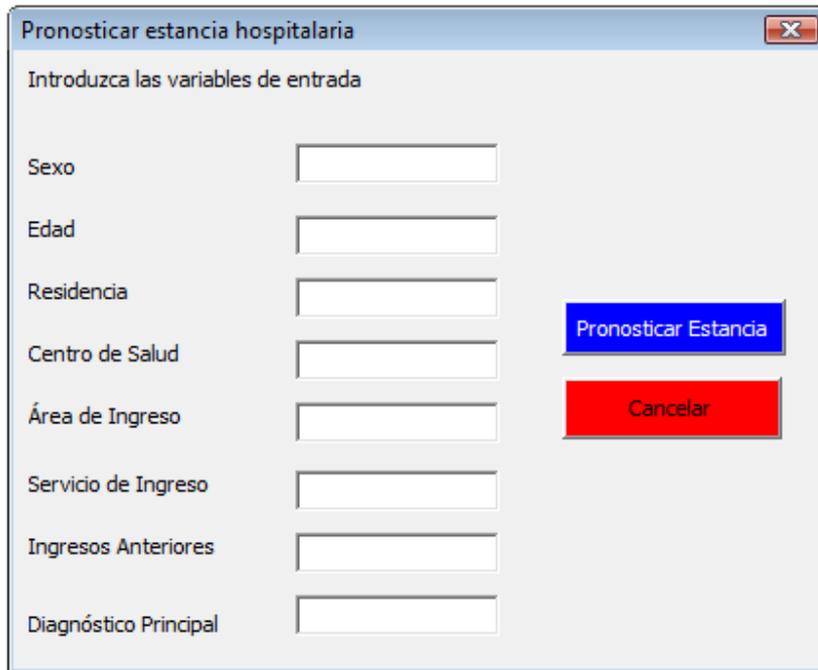
Posteriormente haciendo clic en el ícono Inicio, se puede acceder al menú de módulos, desde donde se puede acceder a cada uno de los 4 módulos, o devolverse a la interfaz de inicio. Este menú de módulos se presenta en la imagen 2.

**Imagen 2. Menú de módulos**



El módulo de pronosticar estancia, presenta una interfaz donde el usuario debe ingresar los datos y características del paciente con cáncer al cual se le quiere proyectar la estancia hospitalaria. La interfaz se presenta en la imagen 3.

**Imagen 3. Interfaz de pronóstico de estancia**



The image shows a software window titled "Pronosticar estancia hospitalaria". Inside the window, there is a prompt "Introduzca las variables de entrada" followed by eight text input fields. The labels for these fields are: "Sexo", "Edad", "Residencia", "Centro de Salud", "Área de Ingreso", "Servicio de Ingreso", "Ingresos Anteriores", and "Diagnóstico Principal". To the right of the input fields, there are two buttons: a blue button labeled "Pronosticar Estancia" and a red button labeled "Cancelar".

El siguiente módulo es el de entrenamiento de la red neuronal, el cual es muy similar al módulo de proyección, con la única diferencia que se debe introducir la salida real, es decir, una vez introducidas las características del usuario con cáncer también se debe indicar la estancia hospitalaria de dicho paciente; esta pantalla se observa en la imagen 4.

**Imagen 4. Interfaz de entrenamiento de la red neuronal**

El siguiente módulo es el de optimización de la red neuronal, accediendo a este módulo se presenta una página de Excel donde se puede observar el valor inicial de los pesos de la red neuronal, así como los resultados en cuanto al error de construcción, el de prueba y el error total de la red; así mismo tiene la opción de ir haciendo cada paso del proceso de optimización, así como iniciar la optimización con 45 iteraciones o agregar 45 iteraciones al proceso. Esta pantalla se presenta en la imagen 5.

**Imagen 5. Pantalla de optimización de la red**

**OPTIMIZAR PESOS W**

Error de prueba	Error de construcción	Error total	F. Aprendiziz	V1							V2				
				01	02	03	04	05	06	07	01	02	03	04	
42.78%	40.80%	42.4%	2.545E-06	-0.790163	2.30544	-1.7766	-1.667	6.36629	-2.3389	-1.6851	1.48246	1.10529	-1.7628	1.031839	
Aleatorio Inicial				0.00049754	-0.066556	-0.1006	0.20807	-0.9223	-1.8743	-1.3534	0.5116	1.97378	-1.1841	-0.1992	0.096786

PRIMERA GENERACION

El cuarto y último modo corresponde a una opción que tiene el instrumento para intervenir en el accionar de la red neuronal, el mismo consiste en la modificación manual de los pesos de la red, así como del factor de aprendizaje. Para esto se accede a una pantalla donde se pueden incluir los valores y luego agregarlo al modelo. La pantalla de acceso se muestra en la imagen 6.

**Imagen 6. Pantalla de ejecución de la modificación manual de los pesos**

### Modificar pesos de la red manualmente

**Instrucciones**

Introduzca los nuevos valores de frecuencia de aprendizaje, pesos de las neuronas de entrada y de las neuronas ocultas. Luego presione el botón azul de la parte inferior de la pantalla.

Frecuencia de aprendizaje		
---------------------------	--	--

Sexo							Edad							Residencia							Centro de Salud						
O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7
W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16	W17	W18	W19	W20	W21	W22	W23	W24	W25	W26	W27	W28	W29

Área de ingreso							Servicio de ingreso							Ingresos anteriores							Diagnóstico principal						
O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7
W30	W31	W32	W33	W34	W35	W36	W37	W38	W39	W40	W41	W42	W43	W44	W45	W46	W47	W48	W49	W50	W51	W52	W53	W54	W55	W56	W57

Neuronas ocultas						
O1	O2	O3	O4	O5	O6	O7
W58	W59	W60	W61	W62	W63	W64

Modificar Factor de Aprendizaje y Pesos

Regresar al Menú

Finalmente, en el menú de módulos, se puede acceder a una consulta directa a los errores de la red, que son los que dictan la eficacia de la misma, este acceso permite conocer en cualquier momento los resultados del último aprendizaje de la red, con el grupo de pesos que están actualmente vigentes, la pantalla que se activa se presenta en la siguiente imagen.

**Imagen 7. Resultados de error de la red**



En esta se puede consultar el error total de la red, el error de construcción y el error de prueba, con lo que se puede validar la eficacia de la misma. Los errores presentados están en términos relativos, al multiplicarlos por cien se obtiene el valor porcentual.

A continuación se presenta una descripción del manejo y la funcionalidad de cada uno de los módulos que componen la herramienta.

## **Módulos**

### **1. Pronosticar estancia hospitalaria**

El módulo de Pronosticar estancia es el que permite hacer proyecciones sobre la cantidad de días que se espera que un paciente con cáncer permanezca hospitalizado. La interfaz de este módulo se presenta en la imagen 8.

**Imagen 8. Interfaz del módulo de pronosticar casos nuevos**

Pronosticar estancia hospitalaria

Introduzca las variables de entrada

Sexo

Edad

Residencia

Centro de Salud

Área de Ingreso

Servicio de Ingreso

Ingresos Anteriores

Diagnóstico Principal

Pronosticar Estancia

Cancelar

Esta interfaz presenta una serie de componentes que se describen a continuación:

### **Variables de entrada**

- 1. Sexo:** Se debe indicar el sexo de cada paciente con cáncer. Se utiliza la siguiente codificación: para el hombre el número 1 y para la mujer el número 2. Estos valores se deben de dividir por 100 para efectos de incluirlos en el módulo indicado en la imagen 8.
- 2. Edad:** Se debe indicar los años de vida del paciente con cáncer en el momento del egreso del hospital de la C.C.S.S. Para efectos de incorporar la edad en la red se establecieron 20 categorías o grupos de edad, con rangos de 5 años cada una. Las categorías tienen valores numéricos del 1 al 20 en orden ascendente según los rangos de edad. Para efectos de incorporar el dato en el módulo visualizado en la imagen 8, se debe dividir el valor por 100.
- 3. Residencia:** Se debe indicar el cantón de residencia en el cual habita el paciente con cáncer al ingresar en uno de los centros de salud de la C.C.S.S. Esta variable

se indica con valores numéricos del 1 al 81, correspondiente al número total de cantones que tiene nuestro país. Estos valores se deben de dividir por 100 para efectos de incluirlos en el módulo indicado en la imagen 8.

- 4. Centro de salud:** Se indica el hospital de la C.C.S.S por el cual ingresó el paciente con cáncer. Para efectos del desarrollo de esta red neuronal se seleccionaron los 12 hospitales que reciben la mayor cantidad de pacientes con cáncer; estos están codificados con valores numéricos del 1 al 12, y para efectos de introducirlos en la herramienta deben dividirse entre 100.
- 5. Área de ingreso:** Se indica el área de atención en el centro de salud por la cual ingresa el paciente con cáncer. Tiene una codificación con números del 1 al 3 (1 para consulta externa, 2 para urgencias y 3 para nacimientos). El valor correspondiente se debe dividir por 100 para efectos de incluirlo en el módulo indicado en la imagen 8.
- 6. Servicio de ingreso:** Se indica el servicio del centro de salud en el cual el paciente con cáncer fue atendido después de haber ingresado. La codificación de esta variable es la siguiente: Medicina (1), Cirugía (2), Gineco-obstetricia (3), Pediatría (4), Cirugía Mayor Ambulatoria (6), Psiquiatría (7), Cuidados intensivos en medicina (8), Cuidados intensivos quirúrgicos (9), Cirugía Vespertina (11). El valor correspondiente se debe dividir por 100 para efectos de incluirlo en el módulo indicado en la imagen 8.
- 7. Ingresos anteriores:** Se indica el número de ocasiones en que un paciente con cáncer ha ingresado a los centros de salud de la C.C.S.S antes del ingreso actual. El valor correspondiente se debe dividir por 100 para efectos de incluirlo en el módulo indicado en la imagen 8.
- 8. Diagnóstico principal:** Se indica el tipo de cáncer principal sobre el cual fue diagnosticado el paciente. Para efectos del desarrollo de esta red neuronal se seleccionaron los 20 diagnósticos que se han presentado con mayor frecuencia en los últimos 10 años. Cada tipo de diagnóstico fue codificado con un valor

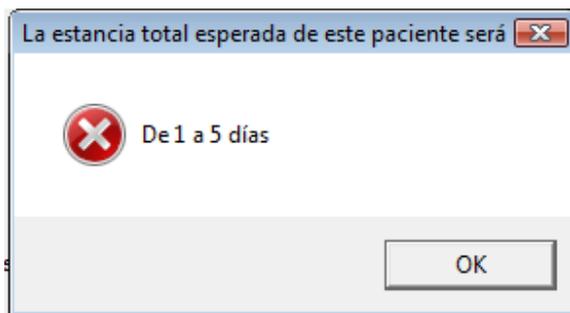
numérico del 1 al 20. El valor correspondiente se debe dividir por 100 para efectos de incluirlo en el módulo indicado en la imagen 8.

**Botón de Pronosticar:** El cual permite ver los resultados de la proyección de estancia hospitalaria según los datos del paciente con cáncer que se indicaron en las variables de entrada.

**Botón de Cancelar:** Este devuelve la herramienta a la pantalla del menú de módulos.

Después de hacer clic en el botón Pronosticar se visualiza la proyección de estancia en una pantalla, tal como se muestra en la imagen 9.

### **Imagen 9. Resultado del pronóstico**



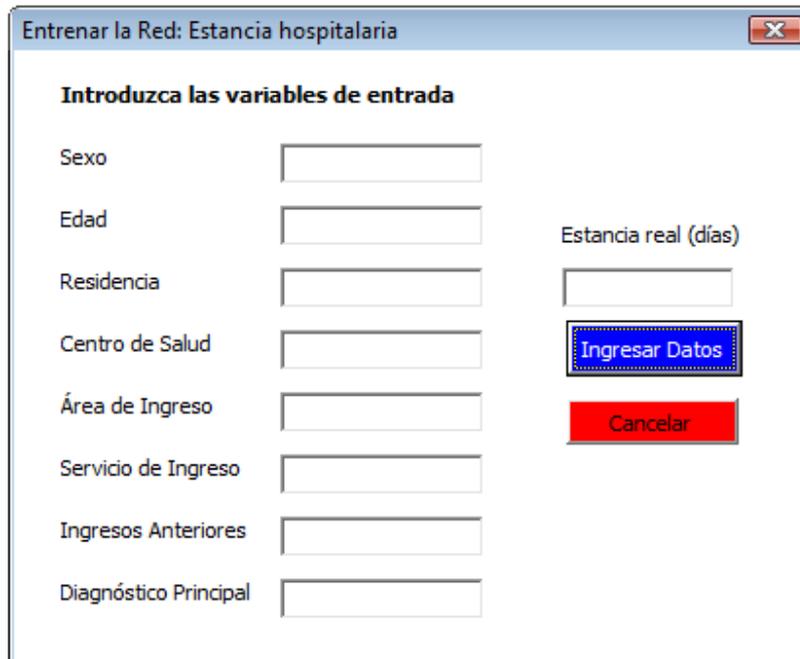
Al hacer clic en el botón OK se regresa al menú de módulos, para continuar viendo otras estimaciones o accediendo a otros módulos.

## **2. Entrenar la Red**

Este modulo permite ejecutar el proceso de entrenamiento de la red neuronal, este proceso consiste en determinar un nuevo conjunto de pesos para las relaciones entre variables, que mejore lo resultados de los pronósticos, se utiliza generalmente cuando, al pasar el tiempo, hay un cambio importante en los datos reales, que puedan afectar los resultados de la red. En este caso debería entrenarse cuando las condiciones hospitalarias o las características de los pacientes con cáncer varíen considerablemente afectando la estancia hospitalaria.

La pantalla del módulo de entrenamiento se presenta en la imagen 10.

**Imagen 10. Interfaz del módulo de Entrenar la Red**



El módulo presenta una interfaz muy similar al módulo de pronósticos.

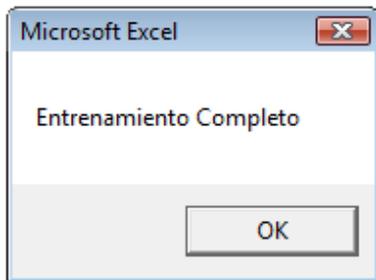
**Variables de entrada:** Deben incluirse en la pantalla las variables de entrada siguiendo las mismas consideraciones establecidas para el módulo de pronósticos

**Estancia real:** Una vez incorporadas las variables de entrada relacionadas con las condiciones y características del paciente con cáncer se indica la estancia hospitalaria real de ese paciente. Para efectos del desarrollo de esta red se crearon 20 categorías de estancia con rangos de 5 días cada una. Esta variable tiene una codificación con números del 1 al 20 (correspondiente a las 20 categorías definidas). El valor correspondiente se debe dividir por 100 para efectos de incluirlo en el módulo indicado en la imagen 10.

**Botón de Ingresar Datos:** Este botón permite ejecutar el entrenamiento de la red. Una vez que se efectúa el entrenamiento, se envía un mensaje de que el proceso fue exitoso

tal como se indica en la imagen 11. Posteriormente se presenta la pantalla de entrenamiento de la red para que se incorporen los datos de otro paciente.

**Imagen 11. Mensaje de entrenamiento exitoso**



**Botón de Cancelar:** Este devuelve la herramienta a la pantalla del menú de módulos.

### **3. Optimizar la red**

El módulo de optimizar la red permite acceder a una pantalla donde se puede hacer el proceso de optimización.

Esta pantalla permite efectuar la optimización de la red de tres formas diferentes, un proceso manual de inicio, compuesto por 18 pasos que conforman 3 iteraciones; un proceso automático que inicia el proceso de optimización con 45 iteraciones y un proceso que agrega 45 iteraciones, pero conservando los datos que se tengan hasta el momento. La pantalla se muestra en la imagen 12, posteriormente se explican cada una de las partes.

**Imagen 12. Pantalla de optimización de la red**

OPTIMIZAR PESOS W														
Error de prueba	Error de construcción	Error total	F. Aprendiz	V1							V2			
				O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4
42.78%	40.80%	42.4%	2.545E-06	-0.790163	2.30544	-1.7766	-1.667	6.36629	-2.3369	-1.6851	1.48246	1.10529	-1.7628	1.031839
Aleatorio Inicial 0.00049754 -0.066556 -0.1006 0.20807 -0.9223 -1.8743 -1.3534 0.5116 1.97378 -1.1841 -0.1992 0.096786 1.0														

Crear población inicial PRIMERA GENERACION

- 1. Botón Limpiar pantalla:** Este botón permite limpiar todos los datos de las diferentes tablas que conforman el proceso de optimización.
- 2. Botón Regresar al Menú:** Este devuelve la herramienta a la pantalla del menú de módulos.
- 3. Botón de Iniciar 45 iteraciones:** Este botón permite hacer un proceso automático de aplicación de 45 iteraciones (rutinas), las cuales iniciar desde cero, es decir que se construyen nuevos individuos (conjuntos de datos) y se van realizando los diferentes procesos para la optimización.
- 4. Botón de Aplicar 45 iteraciones adicionales:** Este botón permite agregar 45 iteraciones de optimización adicionales, que permiten mejorar los resultados que ya se tienen, es decir arranca con un grupo de individuos que ya fueron utilizados anteriormente y sobre esos es que se van mejorando. Este proceso se utiliza normalmente después del de iniciar 45 iteraciones, o después de correr el proceso manual de tres iteraciones de optimización.
- 5. Proceso manual de optimización:** Este proceso consiste en aplicar tres rutinas de optimización, pero haciéndolo paso a paso, mediante 17 botones de color azul que se presentan en la pantalla, los cuales deben ser aplicados en el orden respectivo e inician desde cero, generando una nueva población de análisis (grupo de soluciones) y mejorando a partir de ella.

**6. Información de la red:** La pantalla también tiene un espacio donde se muestra los resultados del error de pronóstico de construcción, prueba y total de la red, según la población que se esté evaluando, esto permite observar mejor el comportamiento de la optimización, para asegurarse que los resultados mejoran el accionar de la red, especialmente en cuanto al error total y el de prueba.

#### 4. Modificar pesos manualmente

Este módulo consiste en una pantalla que permite modificar los pesos de las diferentes conexiones de forma manual, así como la frecuencia de aprendizaje, de manera que un usuario experto, puede hacer ajustes en los pesos, buscando mejorar los resultados. La pantalla se muestra en la imagen 13.

**Imagen 13. Pantalla de modificación manual de pesos**

**Modificar pesos de la red manualmente**

**Instrucciones**

Introduzca los nuevos valores de frecuencia de aprendizaje, pesos de las neuronas de entrada y de las neuronas ocultas. Luego presione el botón azul de la parte inferior de la pantalla.

Frecuencia de aprendizaje	
---------------------------	--

Sexo							Edad							Residencia							Centro de Salud						
O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7
W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16	W17	W18	W19	W20	W21	W22	W23	W24	W25	W26	W27	W28	W29
<input style="width: 100%; height: 100%;" type="text"/>																											

Área de ingreso							Servicio de ingreso							Ingresos anteriores							Diagnóstico principal						
O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7	O1	O2	O3	O4	O5	O6	O7
W30	W31	W32	W33	W34	W35	W36	W37	W38	W39	W40	W41	W42	W43	W44	W45	W46	W47	W48	W49	W50	W51	W52	W53	W54	W55	W56	W57
<input style="width: 100%; height: 100%;" type="text"/>																											

Neuronas ocultas						
O1	O2	O3	O4	O5	O6	O7
W58	W59	W60	W61	W62	W63	W64
<input style="width: 100%; height: 100%;" type="text"/>						

Modificar Factor de Aprendizaje y Pesos

Regresar al Menú

Esta pantalla está compuesta por las siguientes partes:

**1. Área de introducción de pesos:** Está compuesta por 64 espacios para modificar, 56 de conexiones de las variables de entrada con las variables ocultas y 7 de las

ocultas con la de salida, además de un espacio para modificar la frecuencia de aprendizaje.

- 2. Botón Modificar Factor de Aprendizaje y Pesos:** Este botón permite realizar la sustitución de los datos compilados en los espacios, sobre la red neuronal. Cabe mencionar que este proceso modifica todos los datos, lo cual implica que si un dato está en blanco lo sustituirá por cero en el modelo, por lo tanto la modificación de pesos se debe realizar actualizando todos los pesos.
- 3. Botón de Regresar al Menú:** Este devuelve la herramienta a la pantalla del menú de módulos.