

**UNIVERSIDAD NACIONAL  
“SANTIAGO ANTÚNEZ DE MAYOLO”**

**FACULTAD DE CIENCIAS  
ESCUELA PROFESIONAL DE ESTADÍSTICA E INFORMÁTICA**



**ANÁLISIS DE REDES NEURONALES ARTIFICIALES PARA  
EL PRONÓSTICO DE LAS ENFERMEDADES  
PREVALENTES EN EL HOSPITAL VÍCTOR RAMOS  
GUARDIA DE HUARAZ, 2016**

**TESIS PARA OPTAR EL TÍTULO DE:**

**LICENCIADO EN ESTADÍSTICA E INFORMÁTICA**

**Bach. EDWIN ROBER CÓRDOVA DELGADO**

**Bach. JHOJAN PIERRE MAUTINO VIDAURRE**

**ASESOR: DR. ROGER PEDRO NORABUENA FIGUEROA**

**HUARAZ - ANCASH - PERÚ**

**2019**

**Nº Registro: T008**



**FORMATO DE AUTORIZACIÓN PARA PUBLICACIÓN DE TESIS Y TRABAJOS DE INVESTIGACIÓN,  
PARA A OPTAR GRADOS ACADÉMICOS Y TÍTULOS PROFESIONALES EN EL  
REPOSITORIO INSTITUCIONAL DIGITAL - UNASAM**

Conforme al Reglamento del Repositorio Nacional de Trabajos de  
Investigación – RENATI. Resolución del Consejo Directivo de  
SUNEDU N° 033-2016-SUNEDU/CD

**1. Datos del Autor:**

Apellidos y Nombres: **CÓRDOVA DELGADO EDWIN ROBER**

Código de alumno: **04.0258.5.AR**

Teléfono: **997306982**

Correo electrónico: [dany\\_100\\_16@hotmail.com](mailto:dany_100_16@hotmail.com)

DNI O Extranjería: **43002707**

**2. Modalidad de trabajo de investigación:**

Trabajo de investigación

Trabajo académico

Trabajo de suficiencia profesional

Tesis

**3. Título Profesional:**

Bachiller

Título

Segunda especialidad

Licenciado

Magister

Doctor

**4. Título del trabajo de investigación:**

**ANÁLISIS DE REDES NEURONALES ARTIFICIALES PARA EL PRONÓSTICO DE LAS  
ENFERMEDADES PREVALENTES EN EL HOSPITAL VÍCTOR RAMOS GUARDIA DE HUARAZ, 2016**

**5. Facultad de: CIENCIAS**

**6. Escuela, Carrera o Programa: ESTADÍSTICA E INFORMÁTICA**

**7. Asesor:**

Apellidos y Nombres: **NORABUENA FIGUEROA ROGER PEDRO**

Teléfono: **968850116**

Correo electrónico: [rnorabuenaf@unasam.edu.pe](mailto:rnorabuenaf@unasam.edu.pe)

DNI o Extranjería: **41493243**

A través de este medio autorizo a la Universidad Nacional Santiago Antúnez de Mayolo, publicar el trabajo de investigación en formato digital en el Repositorio Institucional Digital, Repositorio Nacional Digital de Acceso Libre (ALICIA) y el Registro Nacional de Trabajos de Investigación (RENATI).

Asimismo, por la presente dejo constancia que los documentos entregados a la UNASAM, versión impresa y digital, son las versiones finales del trabajo sustentado y aprobado por el jurado y son de autoría del suscrito en estricto respeto de la legislación en materia de propiedad intelectual.

Firma: 

D.N.I.:

FECHA:



**FORMATO DE AUTORIZACIÓN PARA PUBLICACIÓN DE TESIS Y TRABAJOS DE INVESTIGACIÓN,  
PARA A OPTAR GRADOS ACADÉMICOS Y TÍTULOS PROFESIONALES EN EL  
REPOSITORIO INSTITUCIONAL DIGITAL - UNASAM**

Conforme al Reglamento del Repositorio Nacional de Trabajos de  
Investigación – RENATI. Resolución del Consejo Directivo de  
SUNEDU N° 033-2016-SUNEDU/CD

**1. Datos del Autor:**

Apellidos y Nombres: **MAUTINO VIDAURRE JOHAN PIERRE**

Código de alumno: **04.0357.9.AR**

Correo electrónico: [mauvi.jp@gmail.com](mailto:mauvi.jp@gmail.com)

Teléfono: **945959091**

DNI O Extranjería: **43738462**

**2. Modalidad de trabajo de investigación:**

Trabajo de investigación

Trabajo académico

Trabajo de suficiencia profesional

Tesis

**3. Título Profesional:**

Bachiller

Título

Segunda especialidad

Licenciado

Magister

Doctor

**4. Título del trabajo de investigación:**

**ANÁLISIS DE REDES NEURONALES ARTIFICIALES PARA EL PRONÓSTICO DE LAS  
ENFERMEDADES PREVALENTES EN EL HOSPITAL VÍCTOR RAMOS GUARDIA DE HUARAZ, 2016**

**5. Facultad de: CIENCIAS**

**6. Escuela, Carrera o Programa: ESTADÍSTICA E INFORMÁTICA**

**7. Asesor:**

Apellidos y Nombres: **NORABUENA FIGUEROA ROGER PEDRO**

Teléfono: **968850116**

Correo electrónico: [rnorabuenaf@unasam.edu.pe](mailto:rnorabuenaf@unasam.edu.pe)

DNI o Extranjería: **41493243**

A través de este medio autorizo a la Universidad Nacional Santiago Antúnez de Mayolo, publicar el trabajo de investigación en formato digital en el Repositorio Institucional Digital, Repositorio Nacional Digital de Acceso Libre (ALICIA) y el Registro Nacional de Trabajos de Investigación (RENATI).

Asimismo, por la presente dejo constancia que los documentos entregados a la UNASAM, versión impresa y digital, son las versiones finales del trabajo sustentado y aprobado por el jurado y son de autoría del suscrito en estricto respeto de la legislación en materia de propiedad intelectual.

Firma: 

D.N.I.:

FECHA:

**HOJA DE VISTO BUENO**



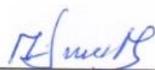
---

MSc. FERNANDO RAÚL ARCE ZÚÑIGA  
PRESIDENTE



---

MSc. JORGE LUIS LLANOS TIZNADO  
SECRETARIO



---

MSc. HUGO WALTER MALDONADO LEYVA  
VOCAL

## **DEDICATORIA**

A Dios, nuestros padres,  
hermanos y familiares quienes  
con su apoyo incondicional  
impulsaron la culminación de  
nuestra carrera profesional.

Edwin y Jhojan

## **AGRADECIMIENTO**

El agradecimiento a la Universidad Nacional Santiago Antúnez de Mayolo, para todos los docentes y autoridades de la escuela académico profesional de Estadística e Informática, que nos orientaron en nuestra vida académica y profesional. Así mismo, el agradecimiento infinito a todos y cada uno de nuestros amigos y compañeros de aula, quienes siempre nos brindaron su aliento permanente y constante para realizar y culminar nuestro trabajo de tesis.

Edwin y Jhojan

# ÍNDICE

DEDICATORIA .....	iii
AGRADECIMIENTO .....	iv
ÍNDICE .....	v
RESUMEN.....	viii
ABSTRACT.....	viii
INTRODUCCIÓN .....	ix
CAPÍTULO I: EL PROBLEMA.....	11
1.1. SELECCIÓN DEL PROBLEMA.....	12
1.2. JUSTIFICACIÓN DEL ESTUDIO .....	14
1.3. OBJETIVOS DE LA INVESTIGACIÓN .....	16
1.3.1. GENERAL.....	16
1.3.2. ESPECÍFICOS.....	16
CAPÍTULO II: MARCO TEÓRICO CONCEPTUAL.....	17
2.1. ANTECEDENTES DEL PROBLEMA.....	18
2.2. PLANTEAMIENTO TEÓRICO ATINGENTE .....	27
2.2.1. Bases teóricas.....	27
2.2.2. Definición de términos .....	76
CAPÍTULO III: HIPÓTESIS Y OPERACIONALIZACIÓN DE VARIABLES .	78
3.1. Hipótesis .....	79
3.2. Operacionalización de variables .....	79
CAPÍTULO IV: METODOLOGÍA .....	80
4.1. Materiales y lugar de ejecución .....	81

4.1.1. Lugar .....	81
4.1.2. Materiales.....	81
4.1.3. Equipos .....	81
4.1.4. Recurso humano .....	81
4.2. Métodos .....	81
CAPÍTULO V: ANÁLISIS Y DISCUSIÓN DE RESULTADOS .....	84
5.1. Análisis de los datos .....	85
5.2. Discusión de resultados .....	113
CAPÍTULO VI: CONCLUSIONES, LIMITACIONES Y RECOMENDACIONES	
CONCLUSIONES .....	117
LIMITACIONES .....	119
RECOMENDACIONES .....	120
REFERENCIAS BIBLIOGRÁFICAS.....	121
ANEXOS .....	125

## RESUMEN

El presente trabajo de tesis según su diseño es no experimental, longitudinal de tipo serie de tiempo y tuvo por objetivo pronosticar las enfermedades prevalentes mediante Redes Neuronales Artificiales en el Hospital Víctor Ramos Guardia de Huaraz, 2016, mediante la aplicación de las redes neuronales artificiales.

Los resultados obtenidos indican que el mejor modelo de pronóstico encontrado es una red neuronal artificial modelo perceptron multicapa con una capa de entrada compuesta por cuatro neuronas, dos capas ocultas compuesta entre 4 a 8 neuronas en la primera capa oculta y entre 4 a 7 neuronas en la segunda capa oculta, y una capa de salida compuesta por una neurona, mediante funciones de activación logística y método de retropropagación como función de corrección del error, así mismo todos los modelos de redes neuronales artificiales encontrados presentan errores globales tendientes a cero.

**Palabras Clave:** Red neuronal artificial, perceptron multicapa.

## **ABSTRACT**

The present work of thesis according to its design is non-experimental, longitudinal series type of time and aimed to predict the prevalent diseases through Artificial Neural Networks in the Hospital Víctor Ramos Guardia de Huaraz, 2016, through the application of artificial neural networks.

The results obtained indicate that the best forecast model found is a multilayer perceptron artificial neural network with an input layer composed of four neurons, two hidden layers composed between 4 to 8 neurons in the first hidden layer and between 4 to 7 neurons in the second hidden layer, and an output layer composed of a neuron, by means of logistic activation functions and backpropagation method as an error correction function, likewise all models of artificial neural networks found have global errors tending to zero.

**Keywords:** Artificial neural network, multilayer perceptron.

## INTRODUCCIÓN

La presente tesis es el resultado de mucha dedicación, esfuerzo y tiempo en la búsqueda de los datos, enlace de bases de datos, la limpieza de las tablas y bases de datos y posterior procesamiento de los datos, para ser analizados mediante las redes neuronales artificiales, para pronosticar.

Con el presente trabajo de tesis se pretende realizar el pronóstico de las enfermedades prevalentes mediante Redes Neuronales Artificiales en el Hospital Víctor Ramos Guardia de Huaraz

Para ello en el capítulo 1, se resalta la descripción del problema de estudio, se muestra en la justificación la importancia que tendrá este trabajo de investigación, y se plantean los objetivos que sirvieron de guías para la realización de la investigación.

En el capítulo 2, se muestran antecedentes, en la que resaltan los estudios de redes neuronales artificiales en estudios referidos al área de la salud. Así mismo, se presenta de manera estructurada el desarrollo de la variable de estudio, con sus respectivas teorías que la fundamentan.

En el capítulo 3, se presenta la hipótesis del presente trabajo de investigación y se realiza la operacionalización de la variable estudiada.

En el capítulo 4, se detallan aspectos metodológicos del presente trabajo de investigación. Así, el lugar de ejecución fue el Hospital Víctor Ramos Guardia de Huaraz, con diseño de investigación de no experimental, de tipo descriptivo, con

una muestra de los registros de atenciones del hospital Víctor Ramos Guardia – Huaraz, comprendidos entre el periodo 2011 al 2016.

En el capítulo 5, se muestran los procedimientos paso a paso para la aplicación de las redes neuronales artificiales, así como el análisis descriptivo respectivo de cada una de las características estudiadas.

Finalmente, en el capítulo 6, se muestra las conclusiones del presente trabajo de investigación y se sugieren recomendaciones que puedan ayudar a una mejor toma de decisiones.

## **CAPÍTULO I: EL PROBLEMA**

## **1.1. SELECCIÓN DEL PROBLEMA**

El hombre en su afán de mejorar sus condiciones de vida, optó por la creación de aparatos médicos tecnificados que consumían demasiado tiempo al producirlas, por lo que, al analizar la vida cotidiana del hombre, esta exigía un control médico a tiempo oportuno, sin que ello signifique que estemos enfermos. En otros casos, sucede lo contrario; nuestro estado general de salud hace que una alteración que podría no considerarse patología, sí se trate como tal.

Para ello hay que tener presente que la patología puede durar un tiempo determinado, durante el cual se ayuda al paciente a salir de este estado de falta de salud. No obstante, la enfermedad puede durar toda la vida, es decir, convertirse en crónica; o puede tener como consecuencia la muerte del individuo, si esta no es tratada a tiempo.

Según la Organización Mundial de Salud (OMS) entre las enfermedades crónicas destaca las cardiopatías, los accidentes cerebrovasculares, el cáncer, las enfermedades respiratorias crónicas y la diabetes. Las discapacidades visuales y la ceguera, los defectos de audición y la sordera, las enfermedades bucodentales y los trastornos genéticos son otras afecciones crónicas que representan una porción sustancial de la carga mundial de morbilidad. Del total de 58 millones de defunciones previstas para 2005 por todas las causas, se estima que 35 millones corresponderán a enfermedades crónicas, lo que supone el doble del número de defunciones correspondiente al conjunto de todas las

enfermedades infecciosas, las enfermedades maternas y perinatales y las carencias nutricionales.

Por tanto, es de vital importancia conocer cómo es que las enfermedades prevalentes irán evolucionando, a fin de tener un pronóstico a futuro y en base a ellas realizar políticas de prevención y agilizar recursos humanos para dar procedimientos oportunos para tratar ésta enfermedad.

Por lo expuesto, se propone analizar las predicciones basadas en Redes neuronales aplicada al diagnóstico médico, que permita conocer el margen de error de los valores futuros de una serie, basada en sus valores pasados, para realizar el correcto entrenamiento de diagnóstico patológico, con la finalidad de reducir el desarrollo de éstas y los decesos de quienes las padecen.

A través de las redes neuronales artificiales (RNA), se desarrollarán las diversas técnicas y herramientas que identificarán a los patrones médicos para diagnosticar de manera óptima las anomalías más frecuentes dentro de una relación causal entre dos series, en este caso los valores pasados de una serie para predecir valores futuros de la serie mediante RNA.

Por consiguiente las redes neuronales artificiales imitan el funcionamiento neuronal biológico, que resolverán problemas complejos en un tiempo corto.

Por lo que las redes neuronales son implementadas en las diversas áreas de salud, la estadística, industria, medicina, seguridad, entre muchas otras,

ayudando de diferente manera con eficiencia y eficacia a resolver los problemas de la organización.

Ante ello, se intenta imitar en forma artificial las habilidades relacionadas con la inteligencia humana. Para lo cual, el uso de ésta técnica, simulará el pensamiento humano mediante conexión de neuronas, que es adecuada sumergirla ante el problema de las patologías diagnosticadas en el Hospital Víctor Ramos Guardia de Huaraz.

Por todo lo mencionado anteriormente, se plantea la siguiente pregunta de investigación:

**¿Será posible pronosticar las enfermedades prevalentes mediante Redes Neuronales Artificiales en el Hospital Víctor Ramos Guardia de Huaraz, 2016?**

## **1.2.JUSTIFICACIÓN DEL ESTUDIO**

El triángulo epidémico, a través de sus tres componentes: Patógeno, Hospedante y Ambiente, constituye la base conceptual en la que se sustenta toda investigación que se desarrolle en el área de la Epidemiología (Bombelli, 2011), donde una de las que posee mayor potencial información para el trabajo es en el área Estadística y Medicina del Hospital, del cual se depende para generar avances tecnológicos que ayuden en la detección temprana de enfermedades, dando el diagnóstico y la prevención oportuna acorde a las políticas de la

institución. Por consecuente, es necesario crear nuevas herramientas con funcionamiento accesible y práctico, que puedan abordar e implementarse en las diferentes áreas de labor médica.

En este estudio de investigación, se buscó una herramienta para citopatólogos y/o residentes del área médica, para que corroboren el diagnóstico de las enfermedades más prevalentes, con el fin de disminuir falsos positivos o falsos negativos. Para lo cual el mencionado triángulo epidémico, se convierte en pirámide al sumarse el hombre como cuarto factor, que interactúa con los tres anteriores para la expresión de la enfermedad. El mismo, que decide estrategias de control de la enfermedad para así aumentar los conocimientos del profesional médico, dentro de las normas establecidas para lograr sistemas de prevención sustentables.

Ante ello los modelos matemáticos y estadísticos, son los que permitirán representar lo que sucede en el diagnóstico de la enfermedad, frente a los distintos patógenos, los cuales se desarrollan dependiendo de los objetivos del trabajo, los datos disponibles, el tipo de experimento y la experiencia del investigador. En general, se utilizan dos estrategias para el modelo de datos epidemiológicos: los modelos empíricos y los mecanísticos. El caso que nos ocupa, utiliza el primero de ellos, como forma de describir el progreso epidémico, a través de la relación observada entre dos o más variables, basándose en algunos o todos los factores capaces de expresar el diagnóstico de la enfermedad, para así tener eficacia y eficiencia en el control de la patología.

Por otro lado, será de gran importancia para los investigadores y profesionales del área estadística, pues dispondrán de información actual de la aplicación de las redes neuronales artificiales.

### **1.3.OBJETIVOS DE LA INVESTIGACIÓN**

#### **1.3.1. GENERAL**

Pronosticar las enfermedades prevalentes mediante Redes Neuronales Artificiales en el Hospital Víctor Ramos Guardia de Huaraz, 2016.

#### **1.3.2. ESPECÍFICOS**

- Identificar las enfermedades más prevalentes en los pacientes del Hospital Víctor Ramos Guardia de Huaraz, 2016.
- Determinar la cantidad de neuronas de entrada para el diseño del modelo de red neuronal artificial para realizar el pronóstico de las enfermedades prevalentes en el Hospital Víctor Ramos Guardia de Huaraz, 2016.
- Determinar el modelo de Redes Neuronales Artificiales adecuado para el pronóstico de las enfermedades prevalentes en el Hospital Víctor Ramos Guardia de Huaraz, 2016.

## **CAPÍTULO II: MARCO TEÓRICO CONCEPTUAL**

## 2.1. ANTECEDENTES DEL PROBLEMA

**Asnate Salazar, Edwin Johnny (Perú-Ancash, 2014).** En su investigación *“Factores asociados al bajo peso de los recién nacidos, del Hospital Víctor Ramos Guardia, Huaraz, 2002-2009, usando minería de datos con redes neuronales”*, el objetivo se centró en determinar los factores de riesgo asociados al bajo peso en los recién nacidos, Hospital Víctor Ramos Guardia, Huaraz, 2002 – 2009, usando minería de datos con redes neuronales, se pudo determinar que los factores maternos más significativos, los factores útero-placentarios de membranas, hemorragia y factor fetal gestación múltiple y por último en la red neuronal perceptron multicapa en la fase de Modelización con la metodología CRIPS-DM que determina los factores de riesgo asociado al bajo peso de recién nacidos, que clasifica de manera correcta a un 94.8475% de las instancias o registros, esta red neuronal es de tres capas, la capa de entrada está compuesta con 29 neuronas, la capa oculta está compuesta por 15 neuronas y la salida con dos.

**Collantes Duarte, Joanna Verónica (Venezuela-Mérida, 2001).** En su investigación *“Predicción con redes neuronales”*, el objetivo fue comparar los métodos de predicción: diseñar, predecir, analizar y comparar los resultados obtenidos con las diferentes metodologías para predicción, utilizando criterios comparativos de errores de ajuste y predicción. En cuanto a las conclusiones las redes neuronales (RN) como herramienta para la predicción han sido de interés recientemente para muchos autores. Una de sus limitaciones para alcanzar un uso generalizado es el establecimiento de la arquitectura de la red.

En este trabajo se proporciona una metodología o conjunto de pasos que pudieran servir de guía para tener éxito en la predicción con las redes neuronales, por lo que la experiencia es necesaria para determinar la arquitectura apropiada de una RN sería muy provechoso la creación de un sistema experto que ayude a identificar las entradas más adecuadas para la red, mediante el uso de los correlogramas simple y parcial, el número de capas ocultas y sus nodos, en el fin, todos los aspectos que involucra el diseño de una RN para predecirla.

**Bernal López, Edwin Andrés (Colombia-Bogotá, 1999).** En su investigación *“Redes Neuronales (Inteligencia Artificial) Aplicada al Diagnóstico Médico”*, el objetivo fue entender la manera adecuada de la exploración que se realizó a través de diferentes aplicaciones de la Inteligencia Artificial en el diagnóstico médico. Por lo que el diagnóstico médico apoyado en la Inteligencia artificial a lo largo de la historia ha presentado una serie de tendencias marcadas por las apariciones de técnicas nuevas en éste campo, instaurando olas que marcan los desarrollos realizados para esta época, tal fue el caso de los sistemas expertos a inicios de los 60. Con las técnicas actuales de la Inteligencia Artificial el campo del diagnóstico médico a abordado problemas más complejos y con una fiabilidad más alta en la obtención de los resultados, bien sea utilizando los métodos clásicos de diagnóstico por probabilidades o los métodos más recientes como los usados en minería de datos. La identificación de un tipo específico de enfermedad es muchas más precisa y sencilla en la actualidad, puesto que muchos de los sistemas de

diagnósticos que se aplican hoy en día, poseen una gran cantidad de conocimiento (plasmado generalmente como bases de datos), además del hecho que este conocimiento es útil para el reconocimiento de enfermedades que presentes características comunes y de ésta manera ampliar la base de conocimiento presente.

**Arana, E.; Marti & Bonmatí (Colombia-Bogotá, 1999).** En su investigación *“Diagnóstico de las lesiones de la calota, selección de variables por redes neuronales y regresión logística”*, el objetivo fue establecer las variables mínimas para diagnosticar las lesiones de la bóveda craneal con tomografía computarizada(TC) y comprobar la precisión de la regresión logística(RL) y las redes neuronales artificiales (RN) en su diagnóstico. Donde se estudió a 167 pacientes con lesiones de bóveda craneal como única manifestación conocida de enfermedad. Los datos clínicos y de la TC se emplearon para los modelos de RL y RN, que se probaron con el método jackknife. Los resultados finales de cada modelo se compararon con el área bajo la curva ROC. Por consecuente los resultados de las lesiones fueron en un 73.1% benignas y 26.9% malignas. No hubo diferencia estadísticamente significativa entre la RL y RN para diagnosticar las lesiones malignas. Para la caracterización de los diagnósticos histológicos, la RN fue estadísticamente superior a la RL. Las variables necesarias para el diagnóstico de lesiones maligna fueron la edad y la definición de los bordes, y para los diagnósticos histológicos, matriz, esclerosis marginal y la edad. Concluyendo que se necesitan cuatro mínimas variables para el diagnóstico de estas lesiones, no siendo importante el tipo de sintomatología.

Las redes neuronales ofrecen grandes posibilidades sobre la estadística para las lesiones de bóveda craneal además de un mejor rendimiento diagnóstico.

**Delgado Alberto (Colombia-Bogotá, 1999).** En su investigación “*Aplicación de las Redes Neuronales Artificiales en Medicina*”, el objetivo fue generar modelos matemáticos que pueden ser entrenados para aprender relaciones no lineales entre un conjunto de datos de entrada y un conjunto de datos de salida, por lo que en Medicina la aplicación más común de estos modelos, es la clasificación de patrones con el propósito de apoyar al profesional médico en el diagnóstico y tratamiento del paciente. Concluyó que las redes neuronales artificiales introducen ventajas de un examen cuantitativo en la práctica médica. Los registros médicos contienen información valiosa que puede ser utilizada para entrenar redes neuronales y crear sistemas expertos, estos enriquecen el diagnóstico del médico general y brindan una nueva perspectiva al médico especialista.

**Nojek & García Martínez (Argentina–Buenos Aires, 2003).** En su investigación “*Modelo de Redes Neuronales optimizadas con algoritmos genéticos, una aplicación para proyecciones sobre pacientes con cáncer hospitalizados en la CCSS*”, con el objetivo disponer de proyecciones efectivas de la demanda de los servicios para realizar una adecuada planificación, garantizando la disposición oportuna de dichos servicios y una utilización eficiente de los recursos. Esta investigación pretende ampliar el conocimiento teórico en relación con la utilización de técnicas de inteligencia

artificial, específicamente las redes neuronales y los algoritmos genéticos, en el desarrollo 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 Caja Costarricense de Seguro Social (CCSS) así como en la estimación de la estancia hospitalaria esperada para este tipo de pacientes. Ante esta premisa de resultados obtenidos como parte del desarrollo de la 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, el cual observamos los resultados en las tres fases. **Fase inicial:** 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%. **Fase**

**Optimización con algoritmos genéticos:** 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. Lo anterior debido a que se lograron reducciones importantes en los porcentajes de error de ambas redes. **Fase de Efectividad del Modelo:** 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 cuatro modelos tradicionales de pronósticos: z 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%. Por tanto, 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 la recomendación es 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. 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. Tal como se determinó a través del trabajo realizado, 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; no obstante, 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 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, como se pudo comprobar en el proceso de investigación realizado, 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.

**Trujilano Cabello, J (España, 2005).** En su investigación “*Redes Neuronales Artificiales en Medicina Intensiva*”, el objetivo fue generar el modelo gravedad del índice de probabilidad de Mortalidad en la Unidad de Cuidados

Intensivos (UCI), en el cual muestra una mala calibración que obliga a buscar modelos reajustados. Puesto que se diseñó, para el reajuste de los modelos, utilizando técnicas basadas en regresión logística (RL) y una red neuronal artificial (RN) (Perceptrón multiplaca con retropropagación del error). Con el fin de valorar los modelos utilizados en un grupo de 964 pacientes que se dividen en un grupo de Desarrollo (736) y un grupo de validación (228), donde se calculó el modelo MPM II-0 y MPM II-24), los modelos reajustados por RL (RLR-0 y RLR-24) y los obtenidos por las RN (RN-0 y RN-24). Por consiguiente los modelos desarrollados se contrastaron con el grupo de validación evaluada en sus propiedades de discriminación en el área, bajo la curva ROC (ABC [IC 95 %]) y su calibración con el test de Hosmer - Lemeshow C (HLC [p]). Resultando los modelos MPM II-0 y MPM-24 obtienen una buena discriminación ( $ABC > 0,8$ ) con pobre calibración ( $HLC > 25$ ) y los modelos reajustados (RLR y RN) mejoran en calibración manteniendo una aceptable discriminación, y la RN es mejor en discriminación ( $ABC = 0,85$  [0,79- 0,90]) y calibración ( $HLC = 21$  [p = 0,005]) en el modelo 24-horas, pero sin alcanzar significación. Por lo que se concluyó que la Red Neuronal Artificial es capaz de estratificar el riesgo de mortalidad hospitalaria utilizando las variables del sistema MPM II. Por lo que en el mismo grupo de pacientes la RN obtienen diferentes probabilidades de morbilidad. Por lo que, esto se asocia con una contribución diferente de las variables en los modelos basados en RL o RN para su posterior diagnóstico preventivo según la patología.

**Guzmán, M. & Carrillo, H (España, 2004).** En su investigación “*Minería de Datos con Redes Neuronales Artificiales: Aplicación en Vacunas – Tuberculosis. Centro de Investigación-Desarrollo y Producción de Vacunas y Sueros. Ave. 27 No. 19805, La Lisa. La Habana. A.P. 16017 Cod. 11600*”.

En su objetivo de hace unos cinco años, un grupo de investigadores del Laboratorio de Dinámica no Lineal de la Universidad Nacional Autónoma de México y del Instituto Finlay de Cuba, se abocaron a explorar las bases digitales de información biomédica aprovechando las nuevas tecnologías para el análisis inteligente de datos y el descubrimiento de conocimiento, salvando así las limitaciones que tiene la aplicación de otros métodos tradicionales a tan grandes volúmenes de datos. Durante este lapso se ha ido desarrollando toda, se observó que no todas tenían la misma incidencia en los 22 años de análisis y el estudio de frecuencias de ocurrencia reveló que, en la década de los 80’s las sustancias en las que se trabajaba más eran los agentes antineoplásticos (45 investigaciones) y la Ciclofosfamida (32 investigaciones), mientras que en la década de los años 90’s otras sustancias, como los adyuvantes inmunológicos, los interferones y los antígenos pasaron a ocupar los primeros lugares (208, 19 116 y 106 respectivamente). Asimismo, se concluyó que actualmente (período 2000-2002), los adyuvantes (128) y los interferones (102) se siguen utilizando pero también se observa una emergente tendencia a la investigación de vacunas sintéticas y de DNA.

## 2.2. PLANTEAMIENTO TEÓRICO ATINGENTE

### 2.2.1. Bases teóricas

#### **Inteligencia Artificial**

**Vásquez, S. (2010)** define la inteligencia artificial (IA) como aquella inteligencia exhibida por artefactos creados por humanos (es decir, artificiales).

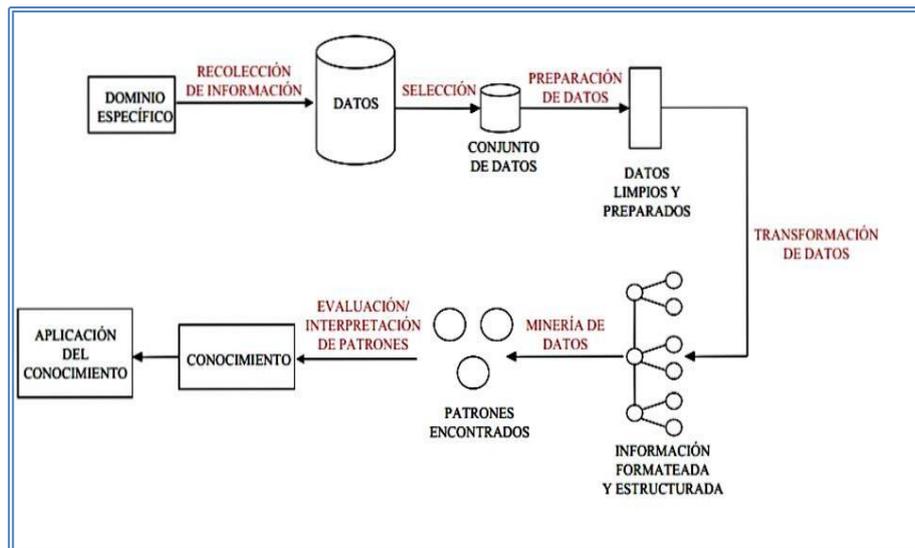
A menudo se aplica a los computadores. El nombre también se usa para referirse al campo de la investigación científica que intenta acercarse a la creación de tales sistemas.

Aunque la IA está rodeada de bastante ciencia ficción, se trata de una rama de la informática, que trata sobre comportamientos inteligentes, aprendizaje y adaptación en máquinas.

#### **Minería de datos (Data Mining)**

**Hernández Orallo, J & Ramirez, F (2004)** en su libro Introducción a la minería de datos. La minería de datos (DM, Data Mining) consiste en la extracción no trivial de información que reside de manera implícita en los datos. Dicha información era previamente desconocida y podrá resultar útil para algún proceso. En otras palabras, la minería de datos prepara, sondea y explora los datos para sacar la información oculta en ellos. Bajo el nombre de minería de datos se engloba todo un conjunto de técnicas encaminadas a la extracción de conocimiento procesable, implícito en las bases de datos. Está fuertemente ligado con la supervisión de procesos industriales ya que resulta muy útil para aprovechar los datos almacenados en las bases de datos. Las bases de la minería de datos se encuentran en la inteligencia artificial y en el análisis estadístico.

Mediante los modelos extraídos utilizando 35 técnicas de minería de datos se aborda la solución a problemas de predicción, clasificación y segmentación. La minería de datos es una fase en el proceso de descubrimiento de conocimiento en base de datos (Knowledge Discovery in Databases, KDD), tal como se presenta en el siguiente diagrama:



Fuente: Hernández Orallo, J & Ramirez, F (2004)

Figura 1: Proceso del descubrimiento de conocimiento de las bases de datos

### Pronostico con Redes Neuronales

**Bails, Dale y Larry Peppers (1993)**, considera que un pronóstico no debe considerarse como algo permanente o estático. Por lo que se debe visualizar como un proceso:

- Determinar el propósito y objetivo del pronóstico.
- Encontrar relaciones teóricas.
- Recolectar la base de datos.
- Analizar la base de datos.

- Estimar un modelo inicial.
- Evaluar el modelo y hacer revisiones.
- Presentar un pronóstico inicial.
- Entrega del pronóstico.
- Establecer procedimientos de monitoreo.

### **Redes neuronales**

El cerebro humano se caracteriza por su gran capacidad de procesar información y su rapidez de respuesta han hecho que se intenten desarrollar sistemas que imiten, al menos en parte sus características.

Las redes neuronales artificiales (RNA) constituye una de las técnicas que intentan reproducir las características del cerebro, a partir de su estructura fisiológica básica: la neurona, así como la agrupación de neuronas en sistemas que pueden mostrar un comportamiento de alguna forma inteligente.

Las RNA son arquitecturas de procesamiento paralelo que ofrecen nuevos mecanismos aplicables a una amplia gama de problemas, permitiendo tratar problemas difíciles o imposibles de resolver mediante métodos tradicionales, su aplicación se ha extendido a problemas de predicción, clasificación, reconocimiento y, en general, a problemas que no pueden ser descritos de forma exacta y para lo que la programación convencional ofrece soluciones limitadas o inadecuadas.

Una definición que enfatiza el aspecto tecnológico, nos dice que las RNA (Artificial Neuronal Networks - ANNs) Son distribuidas (Cuentan con muchas neuronas a través de las cuales distribuyen su memoria), adaptivas (Tienen la capacidad de adaptarse al entorno modificando sus pesos y sinapsis de manera de encontrar una solución aceptable al problema) y generalmente con procesadores elementales (**Pérez, Ma. Luisa & Martín Quintín, 2003**).

Dentro de lo que se entiende por inteligencia artificial (I.A) una de las ramas más prometedoras es la que corresponde a las denominadas redes de autoproceso, entendiendo como tales aquellas formadas por nodos en los que hay elementos procesadores de información de cuyas interacciones locales depende el comportamiento del conjunto del sistema.

Estas redes a su vez pueden ser de procesamiento numérico, la cual recibe directamente la señal de entrada desde el exterior y opera sobre ella. Si la conexión entre los nodos se realiza en forma global bajo unas reglas de composición, estamos ante los llamados sistemas distribuidos, mientras que, si la conexión se realiza agrupando en subredes, se denomina sistema asociativo. Por lo que si se realiza una distribución en la que cada nodo funciona corporativamente, se denomina red neuronal (**Hilera & Martínez, 2000**).

**Pérez, Ma. Luisa & Martín Quintín (2003)** en términos estadísticos, las RNA son estimadores no paramétricos que realizan estimaciones denominadas de

modelo libre. Estas redes se suelen definir en términos de sus algoritmos o implementaciones.

Por consiguiente la Inteligencia Artificial (IA) es un área del conocimiento compuesta por un conjunto de técnicas que se basan en imitar computacionalmente las distintas habilidades relacionadas a la inteligencia del ser humano, como por ejemplo: reconocimiento de patrones, diagnóstico, clasificación, entre otros. Una de estas técnicas imita, específicamente, el comportamiento de las neuronas en el cerebro humano, por lo cual se le ha denominado Redes Neuronales Artificiales (RNA). Es de interés en esta investigación la aplicación de esta metodología para la predicción utilizando series de tiempo.

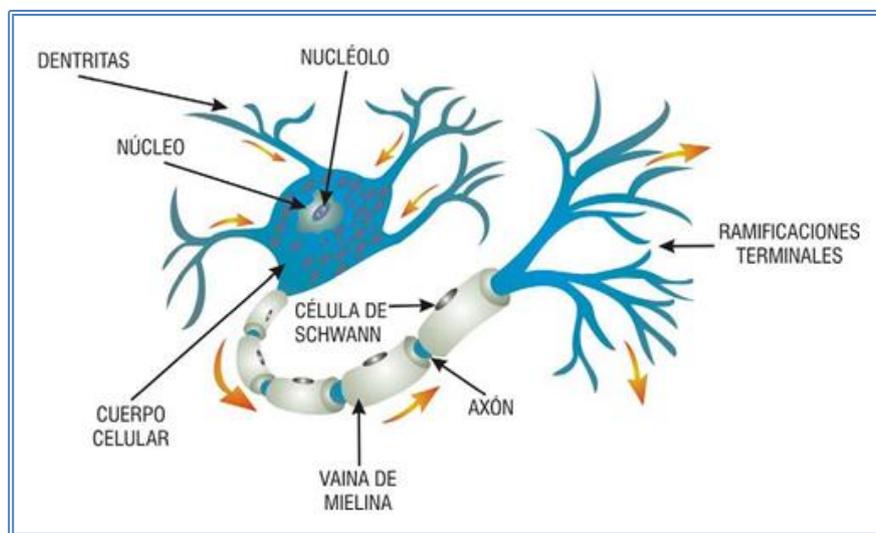
### **Redes Neuronales Biológicas**

Las neuronas son células nerviosas que constituyen los elementos primordiales del sistema nervioso central. Son capaces de recibir señales provenientes de otras neuronas, procesar estas señales, generar pulsos nerviosos, conducir estos pulsos, y transmitirlos a otras neuronas.

La descripción de la estructura y el funcionamiento de los sistemas biológicos nos servirán para comprobar el paralelismo que se intenta establecer en los sistemas artificiales. El sistema nervioso es un sistema complejo, formado por unidades individuales denominadas neuronas, unidas entre sí por una malla de fibras nerviosas.

Las neuronas están separadas estructural, metabólica y funcionalmente y pueden tener tamaños y formas muy variados. Las neuronas forman redes entretejidas muy elaboradas, cuyas funciones varían en diferentes regiones del sistema. Estas diferencias de formas y tamaños reflejan el modo en que las neuronas procesan la información (**Hernández, Miguel, 2013**).

Las neuronas están formadas por tres componentes principales: dendritas, cuerpo celular y axón (**ver figura 2**).



Fuente: Camilo S, (2006)

*Figura 2:* Dibujo Esquemático de Neuronas Biológicas

Las dendritas son como un árbol de redes receptoras de fibras nerviosas que llevan señales eléctricas al cuerpo de la célula (**Hagan, Dermuth & Beale, 1996**).

El cuerpo celular contiene el núcleo, tiene forma esférica y es aquí donde se ejecutan todas las transformaciones necesarias para la vida de la neurona. El axón transmite la señal de salida a otras neuronas. El intercambio químico de

información entre una neurona y otra se hace a través de la sinapsis. Ésta es el punto de interconexión entre neuronas.

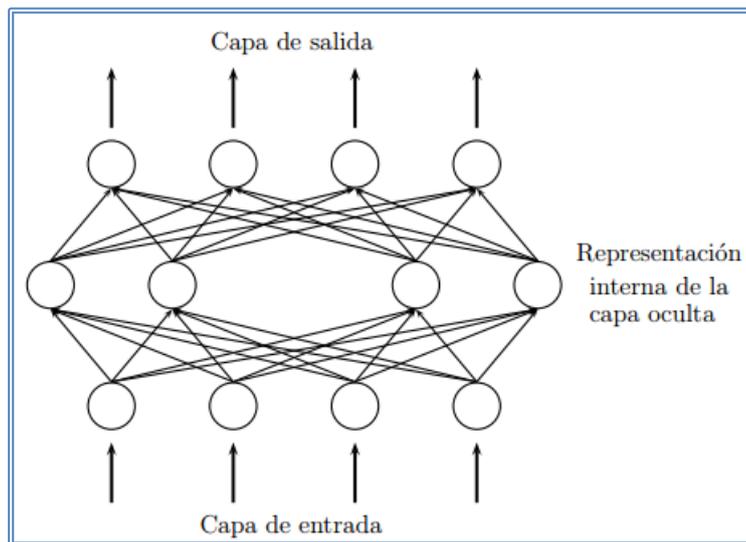
La neurona artificial es una versión muy simple de la neurona biológica, pero puede establecerse fácilmente una equivalencia entre ambas, como lo manifiestan. El cerebro de un ser vivo se modela durante su desarrollo, adquiriendo una serie de cualidades por la influencia de la información que recibe de su entorno. Existen diversas formas de modelar el sistema nervioso: establecer nuevas conexiones, eliminar alguna ya existente, modificar las intensidades sinápticas o eliminar neuronas. La modificación de las intensidades sinápticas es la característica principal que utilizarán las RNA para aprender (Colina & Rivas, 1998).

### **Red Neuronal Artificial**

Una RNA se puede definir como un dispositivo diseñado para la imitación de los sistemas nerviosos de los animales, consiste en una interconexión de unidades, denominadas neuronas artificiales o elementos de proceso, cuyo funcionamiento se inspira en las neuronas biológicas. Simplificando, podemos decir que la función básica de una neurona es sumar sus entradas y producir una salida si la suma es mayor que un umbral determinado. Este tipo de entradas (excitatorio o inhibitorio) y su efectividad, son función de la cantidad de señal recibida en la sinapsis, así como del tipo de estas. Para establecer una similitud directa entre la actividad sináptica y la analogía con las RNA, fijaremos los siguientes aspectos: Las señales que llegan son las entradas a las neuronas; estas

son ponderadas (atenuadas o simplificadas) a través de un parámetro, denominado peso, asociado a la sinapsis correspondiente. Estas señales de entrada pueden excitar a la neurona (sinapsis con peso positivo) o inhibirla (peso negativo). El efecto es la suma de las entradas ponderadas. Si la suma es igual o mayor que el umbral de la neurona entonces la neurona se activa (salida).

Según el investigador **Camilo S, 2006**, las redes neuronales más ampliamente utilizada es la que se conoce con el nombre de Perceptron multicapas, la cual se caracteriza por el hecho de que sus neuronas se agrupan en capas por niveles. Cada una de estas capas está constituida por un conjunto de neuronas. Hay tres tipos de capas diferentes: la capa de entrada, las capas ocultas y la capa de salida, como se observa en la figura 3.



Fuente: Camilo S, (2006)

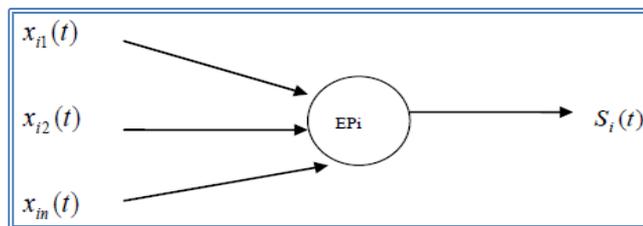
*Figura 3:* Esquema de una Red Neuronal Unidireccional (FeedForward Networks)

Las neuronas de la capa de entrada se encargan únicamente de recibir señales o patrones que vienen del exterior y propagan tales señales a todas las neuronas de la capa siguiente. La última capa actúa como salida de la red, proporcionando al exterior la respuesta de la red para cada uno de los patrones de entrada. Las neuronas de las capas ocultas realizan un procesamiento no lineal de los patrones recibidos. Como se observa en la figura 03, las conexiones del Perceptron multicapas están siempre dirigidas hacia delante, i. e., las neuronas de una capa se conectan con las neuronas de la capa siguiente; por tal motivo reciben el nombre de redes alimentadas hacia delante o redes feedforward. A las conexiones se les asocia un número real llamado peso de la conexión y a las neuronas de la red un umbral, que en el caso del Perceptron multicapas es tratado como una conexión adicional a la neurona (**Camilo S, 2006**).

**Haykin, S. (1994)**. Una red neuronal consiste en un conjunto de unidades de procesamiento, conocidas como nodos o unidades, las cuales están conectadas entre sí. La conectividad de una red neuronal viene dada en términos de una arquitectura o topología, la cual es un grafo con conexiones entre los nodos o unidades. Aquellos nodos que no tienen conexiones de entrada se denominan nodos de entrada y los nodos de los que no sale ninguna conexión se denominan nodos de salida. El resto de nodos se denominan nodos ocultos. Los nodos de computación de una red son los nodos de salida y los nodos ocultos. Todos los nodos que se encuentran a la misma distancia en el grafo de los nodos de entrada forman una capa.

## Unidades de Procesamiento de una Red Neuronal Artificial

**Pérez & Martín (2003)** una RNA, su unidad básica análoga a la neurona biológica, se denomina elementos de proceso de la neurona artificial o simplemente neurona. Distintos modelos de redes utilizan diferentes elementos de proceso, pero se pueden hacer una descripción general de los mismos. Cada elemento de proceso (EP) tiene un conjunto de entradas y una sola salida por las que circulan las señales (Figura 4). Tanto las entradas como la salida dependen del instante de tiempo considerado.



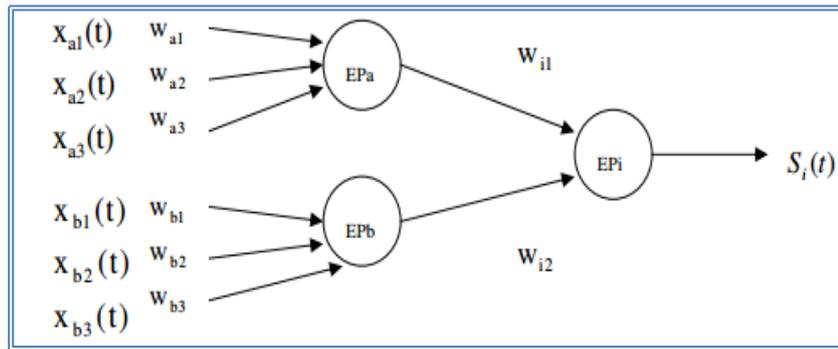
*Fuente: Pérez & Martín (2003)*

*Figura 4:* El elemento de proceso  $i$  con  $n$  entradas y una salida.

Las variables de entrada y salida de una neurona pueden ser discretas o continuas, dependiendo del modelo de neuronas considerado y de la validación que se vaya a dar. Cuando las salidas toman valores continuos, se suele limitar a un intervalo definido, como  $[0, 1]$  o  $[-1, +1]$ .

Las entradas a un elemento de proceso pueden ser las salidas de los elementos de proceso conectados a él, o entradas procedentes del exterior. Así mismo, su salida puede ser una entrada a otros elementos de proceso, o ser una salida de la propia RNA. Cada conexión de entrada tiene asociado un número, denominado peso o fuerza de la conexión, que determina el efecto cuantitativo de unas

unidades sobre otras y corresponde a la sinapsis biológica (figura 5). Los pesos se suelen representar con una  $w$  (weights) y 2 subíndices que indican la neurona y la entrada a dicha neurona a la que están asociados respectivamente. Por tanto, cada neurona tendrá tantos pesos como entradas.



Fuente: Pérez & Martín (2003)

Figura 5: Esquema con varias neuronas conectadas

La entrada total a un elemento de proceso, o entrada neta,  $e_i$  se determina aplicando una regla de propagación,  $\sigma$ , a las entradas y sus correspondientes pesos. Se pueden utilizar diferentes funciones, aunque la más habitual es de tipo lineal y se basa en la suma ponderada de las entradas por los pesos:

$$e_i(t) = \sum_{j=1}^n x_{ij}(t)w_{ij}(t) \quad (2.1)$$

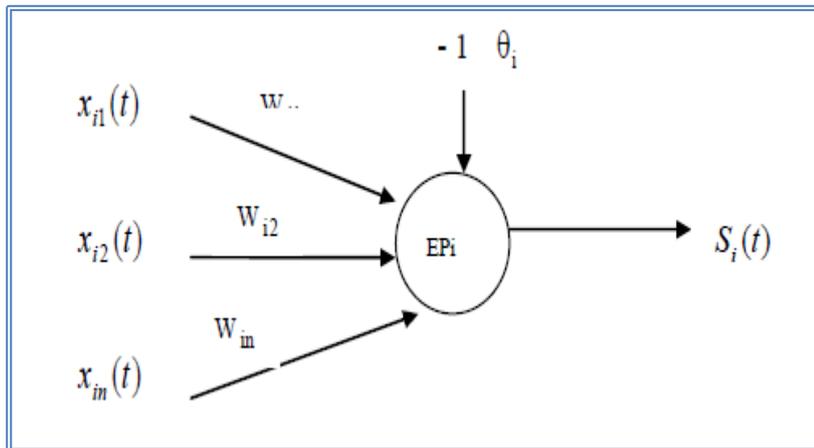
Donde  $i$  representa el elemento de proceso cuya entrada neta se calcula,  $n$  es el elemento de entradas de dicho elemento de proceso, las entradas se representan con una  $x$  y los pesos con  $w$ . Formalmente, también se pueden interpretar como el producto escalar de los vectores de entrada y pesos. En las redes basadas en el cálculo de las distancias entre vectores, se suele utilizar la distancia euclídea, que representa la distancia entre el vector de entrada y el de pesos. Si ambos

vectores son muy similares, la distancia es muy pequeña, aumentado con sus diferencias:

$$e_i(t) = \sqrt{\sum_{j=1}^n [x_{ij}(t) - w_{ij}(t)]^2} \quad (2.2)$$

En algunos modelos la neurona se añade al conjunto de pesos un parámetro adicional,  $\theta_i$ , denominado umbral o tendencia, con lo que se añade un grado de libertad más a la neurona (figura 06). Dicho parámetro se asocia a una entrada con valor fijo igual a -1. La entrada neta calculada en (2.3) será en este caso:

$$e_i(t) = \left[ \sum_{j=1}^n x_{ij}(t)w_{ij}(t) \right] - \theta_i \quad (2.3)$$



*Fuente: Pérez & Martín (2003)*

*Figura 6: Modelo de neurona con umbral*

### **Función de Activación**

La selección de la Función de Activación (FA) depende del criterio del investigador y del problema en estudio, en muchas ocasiones se selecciona por ensayo y error.

Cada elemento de proceso en un instante de tiempo determinado tiene asociado un valor de activación,  $a_i(t)$ . Su nuevo valor de activación se determina aplicando una función,  $F$ , denominada función de activación, al valor de activación previo y la entrada neta.

Al conjunto de valores de activación de los elementos de proceso de una red indica lo que esta representa en un momento dado. Dichos valores pueden ser continuos o discretos.

La función de activación se suele considerar determinista y, en la mayor parte de los modelos, monótona creciente y continua, como se observa habitualmente en las neuronas biológicas. Cuando los algoritmos de aprendizaje requieren de una función de activación sea derivable, las más empleadas son las funciones de tipo sigmoideo.

Tanto la función de salida como la de activación pueden tomar distintas formas, aunque generalmente son funciones de tipo umbral, de modo que un elemento de proceso afecta a otro solo si el nivel de activación del primero supera un cierto valor (**Pérez & Martín, 2003**).

Existen diversos tipos de FA, entre los más utilizados están (**Hilera, J. R. & Martínez V. J., 1995**):

- a) **Función Paso:** La salida de este tipo de FA puede ser 0 o 1, dependiendo si el parámetro de la función es positivo o negativo. Se usa para problemas de clasificación. En forma matemática puede expresarse como:

$$f(v) = \begin{cases} 1 & \text{si } v \geq 0 \\ 0 & \text{si } v < 0 \end{cases} \quad (2.4)$$

- b) **Función Lineal:** La entrada de la FA es igual a la salida. Se usa en diversos tipos de redes, con frecuencia, en la capa de salida. Su expresión matemática es:

$$f(v) = v \quad (2.5)$$

- c) **Función Rampa:** Su salida está entre  $-1$  y  $1$ . Matemáticamente se expresa como:

$$f(v) = \begin{cases} -1 & \text{si } v \leq -1 \\ v & \text{si } -1 < v < 1 \\ 1 & \text{si } v \geq 1 \end{cases} \quad (2.6)$$

- d) **Función Logística o sigmoidea:** Su salida comprende valores entre  $0$  y  $1$ . Es la FA más usada en redes neuronales y se recomienda para problemas de predicción. Su expresión matemática es:

$$f(v) = \frac{1}{1 + e^{-v}} \quad (2.7)$$

- e) **Función Tangente Hiperbólica:** Es semejante a la función logística, pero su salida está entre  $-1$  y  $1$ . Se utiliza con frecuencia en redes multicapas. En forma matemática se expresa como:

$$\Gamma(v) = \text{TanH}(v) \quad (2.8)$$

- f) **Función Gaussiana:** Su rango está entre  $0$  y  $1$ . Se utiliza en redes neuronales de función de base radial, las cuales pueden aplicarse a problemas de predicción. Su expresión matemática es:

$$\Gamma(v) = e^{-v^2/\sigma^2} \quad (2.9)$$

## **Entrenamiento de la red neuronal artificial**

Un modelo de red neuronal al ser construido origina dos actividades fundamentales para la posterior utilización. Estas son: entrenamiento o aprendizaje y generalización o prueba.

El entrenamiento de una red se lleva a cabo mediante una regla o algoritmo de aprendizaje. Este algoritmo es un procedimiento para modificar los pesos de una red y su propósito es entrenar la red para ejecutar alguna tarea (Hagan, Dermuth & Beale, 1996). En otras palabras, los algoritmos de aprendizaje proporcionan la habilidad para aprender a capturar información, mediante el ajuste de los pesos o ponderaciones de interconexión.

En el proceso de entrenamiento se usa cierta terminología que conviene explicar previamente (**Colina & Rivas, 1998**):

- a) **Patrón:** una observación o experimento del conjunto o patrones de entrenamiento.
- b) **Patrones de Entrenamiento:** son ejemplos o grupos de datos suministrados a la red para su aprendizaje y están compuestos por patrones de entrada y patrones de salida.
- c) **Ciclo:** Es un barrido de todos los patrones del conjunto o patrones de entrenamiento.
- d) **Criterios de Parada del Entrenamiento:**
  - Se repiten los ciclos hasta que el error alcance un valor máximo permisible.

- Se alcance un número máximo de ciclos.

### **Algoritmos de Entrenamiento**

Los algoritmos de entrenamiento, en general, pueden clasificarse en dos tipos:

- a) Supervisados:** “un maestro” guía la red en cada etapa del aprendizaje, indicándole el resultado correcto (Colina & Rivas, 1998). La misión del algoritmo es ajustar los pesos de la red de manera tal, que dado un conjunto de entradas las salidas proporcionadas por la red deberán coincidir lo más posible con las salidas especificadas en el patrón de entrenamiento.

Ejemplos: regla de Hebb, aprendizaje de Widrow-Hoff y algoritmo de retropropagación.

- b) No Supervisados:** los pesos son modificados únicamente en respuesta a las entradas de la red, es decir no están disponibles las salidas (u objetivos). La mayoría de estos algoritmos tienen como misión realizar algún tipo de agrupamiento.

Éstos aprenden a categorizar los patrones de entrada en un número finito de clases (Hagan, Dermuth & Beale, 1996). Ejemplos: regla de Hebb no supervisada y regla de Kohonen.

En esta investigación utilizaremos únicamente algoritmos de entrenamiento supervisados. Debido a su adecuación para resolver problemas de predicción. Son muchos los algoritmos de aprendizaje, pero se expondrán

brevemente los que se consideran de interés para la comprensión de este punto y de este trabajo.

c) **Perceptrón Discreto:** se utiliza para clasificar mediante una recta, es un dicotomizador. El algoritmo de aprendizaje para un perceptrón discreto con una sola neurona puede expresarse como sigue:

$$W_{(k+1)} = W_{(k)} + ceX \quad (2.10)$$

Donde:

$W_{(k)}$ : Es el vector de pesos para el patrón de entrenamiento k.

$W_{(k+1)}$ : Es el vector de pesos para el patrón de entrenamiento k+1.

c: Factor de aprendizaje o paso.

e:  $y_d - y_n$ : Es el error, la diferencia entre la salida deseada (proveniente del patrón de entrada del conjunto de entrenamiento) y la salida estimada por la red neuronal.

$X$ : Vector de entradas (proveniente del patrón de entrada del conjunto de entrenamiento).

Según Colina & Rivas, (1998) el método general de entrenamiento puede resumirse en los siguientes pasos:

#### A. Pasos hacia delante:

1. Seleccionar un patrón de entrada del conjunto de entrenamiento.
2. Aplicar esta entrada a la red y calcular la salida.

## **B. Pasos hacia atrás:**

1. Calcular el error entre la salida de la red neuronal y la salida deseada para el patrón de entrada usado.
2. Ajustar los pesos para que el error cometido entre la salida de la red neuronal y la salida deseada sea disminuido.
3. Repetir los pasos 1 al 4 para todos los patrones de entrenamiento, hasta que el error global sea aceptablemente bajo.

Los valores de los pesos se buscan de forma tal que se reduzca la función del error (diferencia cuadrática entre la salida deseada y la salida de la red neuronal), y este proceso se realiza por el método del gradiente descendente. La idea del método es realizar un cambio en los pesos inversamente proporcional a la derivada del error respecto al peso para cada patrón.

Durante la fase de asociación o clasificación, la red opera enteramente en forma de cascada directa o hacia delante, es decir sin retroalimentación. Sin embargo, el ajuste de los pesos obtenido por la regla de entrenamiento se realiza desde atrás hacia delante, pasando por las capas ocultas hasta el nivel de entrada. Como el error se propaga en la red de atrás hacia delante, se le denominó a este procedimiento algoritmo de retropropagación.

### **Error de Entrenamiento (Colina & Rivas, 1998)**

Para medir el error de entrenamiento comúnmente se utiliza la suma cuadrática del error de entrenamiento (SCEE), cuya expresión matemática es:

$$SCEE = (Y_d - Y_n)^2 \quad (2.11)$$

$Y_d$ : Y deseada. Es la salida del modelo, especificada en el patrón de entrenamiento.

$Y_n$ : Y neuronal. Es la salida proporcionada por la red, cuando ésta ya ha sido entrenada y se le proporcionan las entradas del patrón de entrenamiento.

### **Generalización de una Red Neuronal Artificial**

Una vez que una red neuronal ha sido entrenada mediante un algoritmo supervisado, esto implica que los pesos de interconexión han sido ajustados de forma que se minimice el error.

Esta red deberá estar en capacidad de proveer salidas próximas a los valores deseados cuando se le proporcionan nuevos ejemplos, es decir, entradas que no pertenecen al conjunto de entrenamiento, sino que forman parte de los patrones de prueba.

A este proceso se le conoce como generalización de un modelo de RNA.

Según los autores **Colina & Rivas, (1998)**, para que ocurra una buena generalización deben cumplirse las siguientes condiciones:

- Las entradas de la red deben contener suficiente información en relación con el objetivo o salida deseada, de manera que exista una función matemática que relacione correctamente (con el grado de exactitud deseado) las salidas con las entradas.

- La función que se trata de aprender debe ser, en cierto sentido, suave. Es decir, cambios pequeños en las entradas, producen cambios pequeños en las salidas.

El número de patrones de entrenamiento deberá ser suficientemente grande y representativo del conjunto de los casos que se quiere generalizar.

### **Error de generalización (Colina y Rivas, 1998)**

Para medir el error de generalización comúnmente se utiliza la suma cuadrática del error de generalización (SCEG), cuya expresión matemática es:

$$SCEG = (Y_d - Y_n)^2 \quad (2.12)$$

Donde:

$Y_d$ : Y deseada. Es la salida del modelo, especificada en el patrón de prueba.

$Y_n$ : Y neuronal. Es la salida proporcionada por la red, cuando ésta ha sido entrenada previamente y se le proporcionan las entradas del patrón de prueba

### **Modelos de redes neuronales usados para predicción**

Son muchos los autores que han intentado hacer predicciones sobre los valores futuros de una variable mediante redes neuronales, nombraremos algunos que se consideran de interés y se explicará brevemente el modelo de red que utilizaron:

**Faraway y Chatfield (1998)** exponen un caso de estudio ajustando una variedad de modelos de redes neuronales a los bien conocidos datos de la línea aérea y

compara los resultados de la predicción con los obtenidos con el método de Box-Jenkins y Holt-Winters.

Se consideran RN de alimentación adelantada, una capa oculta con función de activación logística, las entradas corresponden a distintos retrasos de la variable y una salida con función de activación lineal que es la predicción en el tiempo  $t$ . la figura 07, representa un ejemplo de una red neuronal para predicción, con tres entradas: una entrada con valor constante 1, el valor de la variable en el tiempo  $t-1$ , y el valor de la variable en el tiempo  $t-12$ , una capa oculta con dos nodos o neuronas y una sola salida:

#### Procedimiento

- 1) Se calcula  $v_j = \sum w_{ij}y_i$  donde los  $w_{ij}$  denotan los pesos de las conexiones entre la entrada  $y_i$  y la  $j$ -ésima neurona.

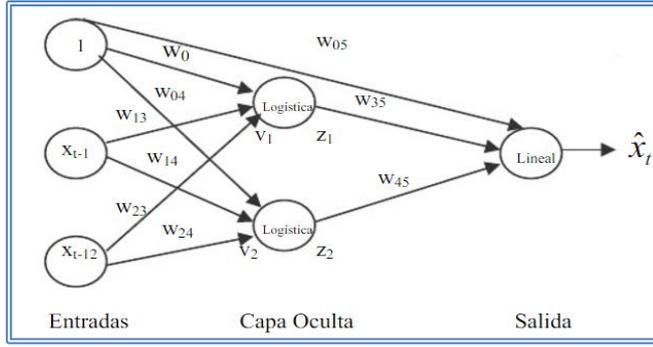
Los valores de entrada, en este caso, son

$$y_1 = 1, y_2 = x_{t-1}, y_3 = x_{t-12}.$$

- 2) La suma lineal,  $v_j$ , es transformada por la función de activación logística:

$$z_j = \frac{1}{(1+\exp(-v_j))}, \quad \text{para } j=1,2; \text{ la cual arroja valores entre 0 y 1.}$$

- 3) Una operación similar se aplica a  $z_1$ ,  $z_2$  y a la constante de entrada para obtener la salida o predicción. Se aplica una función lineal.



Fuente: Faraway & Chatfield (1998)

Figura 7: RN de alimentación adelantada con una capa oculta para la predicción de series de tiempo con datos mensuales.

Ecuación de Predicción General:

$$\hat{x}_t = \phi\{w_{co} + \sum_h w_{ho}\phi_h(w_{ch} + \sum_i w_{ih}x_{t-j})\} \quad (2.13)$$

Donde  $\{w_{ch}\}$  denota los pesos para las conexiones entre la entrada constante y la neurona oculta y  $w_{co}$  denota el peso de la conexión directa entre la entrada constante y la salida. Los pesos  $\{w_{ih}\}$  y  $\{w_{ho}\}$  denotan los pesos para las otras conexiones entre las entradas y las neuronas ocultas, y entre estas neuronas y la salida, respectivamente. Las dos funciones  $\phi_h$  y  $\phi_o$  denotan las funciones de activación usadas en la capa oculta y en la salida respectivamente. Los pesos a ser usados en el modelo de RN son estimados de los datos minimizando la suma de cuadrados de los errores de predicción,

$$s = \sum_t (\hat{x}_t - x_t)^2 \quad (2.14)$$

Sobre la primera parte de la serie de tiempo, llamada el conjunto de entrenamiento (los pesos iniciales se generan aleatoriamente 50 veces y se toma el modelo que tenga S mínimo). Se utiliza el algoritmo de retropropagación.

**Hill, T. O'Connor & Remus, W. (1996):** El experimento consistió en comparar las predicciones producidas mediante RNA con las predicciones obtenidas por expertos, utilizando seis métodos estadísticos de series de tiempo, en la famosa competencia de **Makridakis, S** en 1982. Una arquitectura de RNA fue implementada para las series anuales, otra para las series trimestrales y otra para series mensuales. En todos los casos, la RNA es de alimentación adelantada, totalmente conectada, tiene una sola capa oculta con función de activación logística y se usa el algoritmo de retropropagación para su entrenamiento.

Para datos anuales, se consideraron tres entradas, dos nodos en la capa oculta y una salida, es decir, los datos de los años  $t-2$  hasta  $t$  fueron usados para predecir el año  $t+1$ . Para datos trimestrales, se usaron cuatro entradas, dos nodos en la capa oculta y una salida; los datos de los trimestres  $t-3$  hasta  $t$  fueron usados para predecir el trimestre  $t+1$ . Para los datos mensuales, se consideraron nueve entradas, cuatro nodos en la capa oculta y una salida; los datos de los meses  $t-8$  hasta  $t$  fueron usados para predecir el mes  $t+1$ . Para hacer predicciones más allá del periodo  $t+1$ , se obtiene primero la predicción  $t+1$  y se usa como una de las entradas a la red para producir la predicción para el periodo  $t+2$ . Este procedimiento se repite para todo el horizonte de predicción.

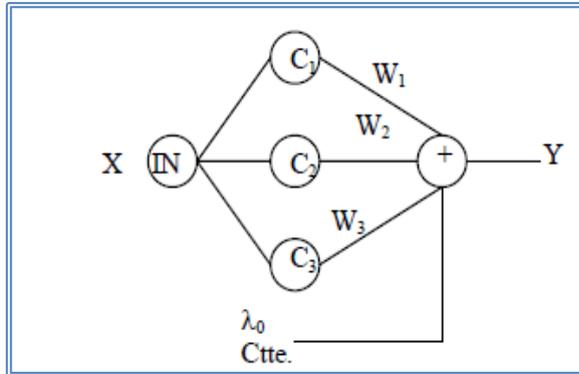
**Wong, F (1991)** también utilizó para la predicción con series de tiempo y redes neuronales de alimentación adelantada, complemente conectadas, con una o dos capas intermedias, una o más salidas y el algoritmo de retropropagación para entrenarlas. Como regla utilizó en la primera capa oculta el doble de nodos que el número de entradas, la mitad de los nodos con función de activación logística y la otra mitad sinusoidal. En la segunda capa utilizó únicamente la función logística.

**Wedding & Cios (1996)** utiliza una RNA denominada Función de Base Radial (FBR) para predicción con series de tiempo. Esta red está compuesta por tres capas.

La primera es la de entrada, alimenta los datos de entrada a cada uno de los nodos de la segunda capa o capa oculta. La segunda capa difiere de las anteriores en que cada nodo representa un grupo de datos, el cual es centrado en un punto particular y tiene un radio dado. La tercera capa o capa de salida consiste de un solo nodo. Esta suma las salidas de los nodos de la segunda capa para alcanzar el valor de decisión.

Cuando un vector de entrada  $x$  es alimentado, en cada nodo de la capa oculta se calcula la distancia del vector de entrada al centro del nodo  $c$ . El valor de la distancia es transformado mediante una función gaussiana y el resultado es la salida del nodo. Ese valor es multiplicado por una constante o valor ponderado. El producto es alimentado en el nodo de la tercera capa, el cual suma todos los productos y cualquier entrada numérica constante, en este caso, se usa la media de todos los datos de entrada. Por último, la salida del nodo de la tercera capa

proporciona el valor de decisión. Representación gráfica de una RN RBF con 3 centros de datos (Colina y Rivas, 1998):



Fuente: Wedding & Cios (1996)

Figura 8: Red Neuronal Función de Base Radial con 3 centros de datos

En forma matemática la RN recibe un vector de entrada  $k$  dimensional y sale un valor escalar usando la siguiente fórmula:

$$f_x(\bar{x}) = \lambda_0 + \sum_{i=1}^n \lambda_i \phi(v_i) \quad (2.15)$$

Donde  $n$  es el nº de nodos de la capa oculta,  $\lambda_i$  son los pesos y  $v$  es el vector de longitud  $n$  que contiene la distancia medida del vector de entrada  $x$  a cada uno de los centros.

Usualmente, la Norma Euclideana es usada para calcular la distancia, pero cualquier otra métrica puede ser usada. La ecuación de la Norma Euclideana es dada por:

$$v_i = \sqrt{\sum_{j=1}^n (x_j - c_{ji})^2} \quad (2.16)$$

$c$ : Centro del grupo (los centros se seleccionan del vector de aprendizaje usando mínimos cuadrados ortogonales).

$\emptyset(v)$  Es la función Gaussiana y está dada por:

$$\emptyset(v) = \exp\left(-v^2/\beta^2\right) \quad (2.17)$$

Donde  $\beta$  es el radio del grupo (se fija en 0,1)

Los datos de entrenamiento y prueba son normalizados y sus predicciones tienen un sesgo añadido: la media.

### **Una Metodología para predicción con Redes Neuronales (Colina y Rivas, 1998)**

En esta parte de la investigación, se definirá una metodología propia para la aplicación de redes neuronales con la cual se realizarán las predicciones de series de tiempo univariable y las predicciones de una serie de tiempo que tiene una relación causal con otra (caso bivivariable).

### **Metodología para predicción de un modelo univariable con redes neuronales (Faraway & Chatfield (1998) & Hill, O'Connor & Remus, 1996)**

A continuación, se especifican los pasos a seguir para construir, entrenar y probar una red neuronal para predecir valores futuros de una serie de tiempo, basada únicamente en sus valores pasados, según los autores:

### **A. Escalamiento de los datos.**

Transformar los datos a valores comprendidos entre 0 y 1, utilizando la siguiente fórmula:

$$z_t = \frac{y_t - Min}{Max - Min} \quad (2.18)$$

Donde:

$y_t$ : Son los valores originales de la serie de tiempo.

$Min$ : Valor mínimo de la serie de tiempo.

$Max$ : Valor máximo de la serie de tiempo.

$z_t$ : Serie de tiempo transformada en valores entre 0 y 1.

### **B. Patrones de entrenamiento y prueba.**

Los valores de la serie de tiempo se dividen en dos conjuntos de datos:

- Patrones de entrenamiento: Está formado por el 80 % de los datos de la serie. Se seleccionan en forma consecutiva y ordenada. Este conjunto de datos es el que se utilizará para el entrenamiento de la red neuronal.
- Patrones de prueba: Está formado por el 20% de los datos de la serie. Corresponden a los datos restantes, una vez que se han seleccionado los patrones de entrenamiento. Este conjunto de datos se utiliza para evaluar la capacidad de generalización o predicción de la red.

### **C. Topología de la red neuronal.**

- Dirección de la información: Alimentación adelantada
- Tipo de interconexión: Totalmente conectada
- Nº de entradas:  $p + 1$  (una constante de valor 1, denominada sesgo o intercepto).

- N° de capas ocultas: 1
- N° de nodos en la capa oculta: q
- N° de salidas: 1
- Función de activación de los nodos de la capa oculta: logística
- Función de activación de la salida: logística.

#### **D. Determinación de las entradas (p).**

Pueden considerarse varias recomendaciones que ayudarán en la selección de las entradas a las REDES NEURONALES ARTIFICIALES:

- La periodicidad de los datos: Como en esta investigación se utilizan únicamente series de tiempo mensuales, es lógico pensar en considerar 12 o 13 retrasos. Pudiera ser conveniente construir una primera red con 13 entradas, correspondientes a los 13 retrasos y analizar los pesos, de manera que se seleccionen las entradas asociadas a los pesos de mayor magnitud, como lo sugieren **Faraway y Chatfield (1998)**. Este análisis de los pesos ayuda a identificar las variables de entrada más importantes.

Una vez determinado el modelo ARIMA, seleccionar como entradas a los valores correspondientes a los retrasos de  $y_i$  involucrados en este modelo.

Otra herramienta que podría considerarse son los correlogramas simple y parcial de la serie estacionaria. Las entradas a la red serían los datos correspondientes a los retrasos que resultaran con una correlación significativamente diferente de cero.

Pruebas por ensayo y error: Otras pruebas pudieran ser de interés, por ejemplo, considerar los datos retrasados 1,4,8,12 periodos, 1,3,6,9,12 periodos, o 1,2,3,4 periodos.

#### **E. Determinación del número de nodos de la capa oculta ( $q$ ).**

Una regla ad hoc, que en experimentos previos ha resultado de utilidad, es asumir que el valor inicial del número de nodos de la capa oculta sea igual al promedio entre el número de entradas y salidas, es decir:  $(\# \text{ Entradas} + \# \text{ Salidas})/2$  (si el valor obtenido es decimal se redondea).

Pueden realizarse pruebas por ensayo y error, agregando más nodos, y comparando los errores de ajuste y predicción.

#### **F. Algoritmo de entrenamiento: Retropropagación.**

#### **G. Selección de los pesos iniciales**

La escogencia de los pesos iniciales puede ser crucial y es recomendable probar con diferentes conjuntos de valores iniciales para tratar de obtener buenos resultados. Los pesos iniciales se generan aleatoriamente 50 veces (Faraway y Chatfield, 1998). Se selecciona el modelo que obtenga el menor promedio entre la suma de cuadrados de los errores de ajuste y predicción (ecuación 2.18).

#### **H. Entrenamiento de la red neuronal seleccionada**

Para entrenar la red es necesario establecer los siguientes parámetros:

- El número máximo de ciclos y el error permitido de convergencia se fijará por ensayo y error.

- Tasa de aprendizaje, incremento de la tasa de aprendizaje y momento pueden ser fijados en 0,05; 1,05 y 0,95, respectivamente. Es conveniente realizar pruebas cambiando estos valores, y evaluando el comportamiento de los errores de entrenamiento y generalización.
- Una vez definida la RNA, con su ecuación se generan los valores de la serie de tiempo ajustada o producida por la red, utilizando los patrones de entrenamiento.
- Se calcula el error de entrenamiento.

### **I. Predicción**

- Usando la ecuación de predicción definida por la RNA se obtiene el valor de predicción  $t+1$ . Para hacer predicciones más allá del periodo  $t+1$ , se utiliza ésta como entrada para producir la predicción  $t+2$  y así sucesivamente para todo el conjunto de predicción.
- Cálculo del error de generalización.

### **J. Comparación entre red neuronal artificial con diferentes entradas**

Mediante el error de entrenamiento y de generalización se comparan las RNA generadas y se selecciona aquella en la que ambos valores sean mínimos. No es conveniente que el error de entrenamiento sea muy pequeño en comparación con el error de generalización, pues esto indica un sobreajuste o memorización. La correlación entre los valores originales de la serie y los estimados por la RNA puede usarse como una medida de la exactitud de la predicción.

**Metodología para predicción de un modelo bivariable con redes neuronales  
(Faraway & Chatfield (1998) & Hill, O'Connor & Remus, 1996)**

La metodología a seguir en el caso bivariable es esencialmente la misma que en el caso univariable, la diferencia radica en las entradas de la red. No sólo se deben considerar retrasos de la serie a predecir (y), sino también valores presentes y pasados de la serie relacionada (x). Para definir las entradas pueden considerarse los siguientes aspectos:

- Periodicidad de los datos: Si se consideran datos mensuales, es lógico pensar en un máximo de 12 o 13 retrasos por serie.
- Los retrasos de interés para las series de entrada y salida pueden identificarse, como en el caso univariable, mediante el modelo o los correlogramas simple y parcial (se seleccionan los retrasos cuyas correlaciones sean significativamente diferentes de cero).
- Modelo de Función de Transferencia: Pudieran considerarse como entrada los valores de la serie de entrada (x) correspondientes a los b retrasos, indicados en el MFT. También los valores de la serie de salida (y) correspondientes a los r retrasos.
- Correlación Cruzada: Una herramienta disponible es considerar como entradas los valores correspondientes a los retrasos de la serie x que tengan correlaciones cruzadas significativamente diferentes de cero.
- Pruebas por ensayo y error: Otras pruebas con diferentes retrasos para ambas series pudieran ser de interés.

Se selecciona la RNA con los menores errores de entrenamiento y generalización.

### **Epidemiología**

La epidemiología es, en la aceptación más común, el “estudio de las epidemias” es decir, de las “enfermedades que afectan transitoriamente a muchas personas en un sitio determinado”. Su significado deriva del griego Epi (sobre) Demos (pueblo) Logos (ciencia). Una definición técnica es la que propone que la epidemiología es “el estudio de la distribución y determinantes de enfermedades en poblaciones humanas”. La epidemiología investiga la distribución, frecuencia y determinantes de las condiciones de salud en las poblaciones humanas, así como las modalidades y el impacto de las respuestas sociales necesarias para atender estas necesidades. La distribución, frecuencia y determinantes de la enfermedad (**Salud Pública y Epidemiológica, 2014**).

### **Morbilidad**

La morbilidad se refiere a los efectos de una enfermedad en una población en el sentido de la proporción de personas que la padecen en un sitio y tiempo determinado. En el sentido de la epidemiología se puede ampliar al estudio y cuantificación de la presencia y efectos de una enfermedad en una población.

También es una adaptación mal empleada al español que proviene de la inglesa “morbidity” generalmente usada así en Sudamérica para identificar una condición médica en la que se estudia los padecimientos de una enfermedad y cuyo verdadero significado es “patológico”. O también para definir

discapacidad, es el porcentaje de individuos que contrae una cierta enfermedad en una población (**Collin, L., 1989**).

### **Prevalencia**

La prevalencia es frecuencia estadística (absoluta relativa) con la que se presenta un determinado fenómeno en una población (entendido este término en su sentido estadístico). En epidemiología los fenómenos a los que suele aplicarse este concepto son enfermedades y la población es humana con la peculiaridad y las derivadas dadas por su aplicación a conjunto de seres humanos y enfermedades o fenómenos relacionados con la salud.

La prevalencia se define como el número de casos de una enfermedad o evento en una población en un momento dado. Existen dos tipos de prevalencia: prevalencia puntual y prevalencia de periodo.

#### **A. Prevalencia puntual o instantánea**

Denominada también tasa de prevalencia, proporción de prevalencia o casos prevalentes. Se habla de prevalencia de punto, puntual o instantánea para referirse a dos cosas:

- Al número o a la frecuencia absoluta de personas que presenta una cierta característica (normalmente una enfermedad) en el instante dado.
- A la proporción o a la frecuencia relativa de individuos de la población que presenta dicha característica.

La prevalencia puntual se estima con la siguiente fórmula:

$$\textit{Prevalencia puntual} = \frac{\textit{Número de casos o antiguos}}{\textit{Población de estudio}}$$

Es la medida estimada en las llamadas encuestas de prevalencia o transversales.

## B. Prevalencia de periodo

Algunos autores la denominan prevalencia periódica o prevalencia lápsica. La prevalencia de periodo se refiere al número o la proporción del total de población que habría presentado ese fenómeno o característica en alguna selección se ha realizado en base a las normas internacionales, que permite priorizar las enfermedades en base a los siguientes criterios:

- **Magnitud:** tiene alta incidencia o prevalencia.
- **Trascendencia:** tienen importante letalidad y/o repercusión social y económica.
- **Vulnerabilidad:** pueden modificarse a través de intervenciones de los servicios de salud.
- **Costo:** tienen costo para los servicios de salud y la sociedad en su conjunto.
- **Vigilancia Internacional:** están sujetas a reglamentos sanitarios internacionales.

La notificación debe realizarse en forma inmediata general, en dependencia del potencial epidémico de las enfermedades y las condiciones de control que requieran. Son responsable de la notificación de todo el personal de salud de los establecimientos pertenecientes a los subsectores: público, seguros sociales,

iglesias, ONG'S y privados en los distintos niveles de atención, de acuerdo a las funciones establecidas.

### **Enfermedades de notificación inmediata**

Las enfermedades y eventos de notificación inmediata, constituyen enfermedades o daños con alto potencial epidémico, que están sujetas a vigilancia internacional o que requieren de actividades de control epidemiológico rápido. En este grupo de enfermedades o daños se pueden resultar de riesgos epidémico, o que, si bien tiene registros de casos, su aparición no es frecuente.

Todo caso sospechoso de estas enfermedades o daños debe ser notificado inmediatamente a los niveles superiores establecidos en el sistema. Una cifra acumulada de estos casos se reporta en forma semanal.

Además de las enfermedades todo evento agudo o borde debe ser comunicado bajo el mismo procedimiento, para contemplar posteriormente la información recolectada.

En ellas las dispuestas por el ministerio se encuentran: sarampión, rubéola, tos, ferina, difteria, parálisis flácida aguda, fiebre amarilla, dengue hemorrágico, dengue clásico, fiebre hemorrágica peruana, peste, meningocócica, cólera, enfermedades por hanta virus, rabia humana.

### **Enfermedades de notificación semanal**

Son patológicas de condición incluidas en la categoría anterior, léxico a que por su potencial epidémico o carácter endémico, no requieren de una notificación inmediata ya que no se realizan actividades especiales de control epidemiológico rápido, permitiendo obtener información que contribuya a la adecuación de los servicios en el enfrentamiento apropiado durante fases epidémicas. Entre ellas se encuentran Tétanos neonatal, hepatitis viral, parotiditis, úlcera genital, flujo uretral, VIH/SIDA, enfermedad diarreica aguda, IRA sin neumonía, neumonía malaria, leishmaniosis, entre otras.

### **Fuentes de notificación**

No existe un sistema nacional de salud al subsector público, registros sociales y organizaciones no gubernamentales, se constituyen en fuentes fundamentales de notificación, los mismos están organizados en espacios geográficos de acuerdo a capacidad resolutiva las Redes de Servicios de Salud (**Salud Pública y Epidemiológica, 2014**).

### **Clasificación Internacional de Enfermedades 10(CIE 10)**

#### **A. Definición**

La clasificación de enfermedades se le conoce como un sistema de categorías a las cuales se les asignan morbosas de acuerdo con criterios establecidos.

## **B. Propósito**

Permitir el registro sistemático e análisis, la interpretación, comparación de datos de mortalidad y morbilidad recolectados en diferentes países o áreas y en diferentes épocas.

## **C. Utilidad**

Convertir los términos diagnósticos de datos problemáticos de salud, de palabras a códigos alfanuméricos que permiten su fácil almacenamiento y posterior recuperación para el análisis de la información. En la práctica se ha convertido en una clasificación diagnóstica estándar mencionada para todos los propósitos epidemiológicos generales y muchos otros de administración de salud.

El CIE, no es adecuada para indagar entidades técnicas individuales y es adecuada para clasificar enfermedades y otros tipos de problemas de salud, consignados en distintos tipos de registros vitales y de salud.

## **D. Estructura básica de la CIE 10° Revisión**

La CIE es un sistema de clasificación de ejes variables cuyo esquema debe servir a todos los propósitos prácticos y epidemiológicos. Este patrón puede ser identificado en los capítulos de la CIE y hasta el momento es considerado como la estructura más útil que cualquiera de las alternativas que se han probado.

La CIE fue publicada por la Organización Mundial de la Salud. Se utiliza a nivel internacional para fines estadísticos relacionados con morbilidad y mortalidad, los sistemas de reintegro y soportes de decisión automática en

medicina. Este sistema está diseñado para promover la comparación internacional de la recolección, procesamiento, clasificación y presentación de estas estadísticas. La CIE es la clasificación central de la WHO Family of International Classifications (WHO-FIC) (en español, la Familia de Clasificaciones Internacionales de la OMS).

La lista CIE-10 tiene su origen en la «Lista de causas de muerte», cuya primera edición editó el Instituto Internacional de Estadística en 1893. La OMS se hizo cargo de la misma en 1948, en la sexta edición, la primera en incluir también causas de morbilidad. A la fecha, la lista en vigor es la décima, y la OMS sigue trabajando en ella.

La CIE-10 se desarrolló en 1992 y su propósito fue rastrear estadísticas de mortalidad. La OMS publica actualizaciones menores anuales y actualizaciones mayores cada tres años. Posteriormente, algunos países han creado sus propias extensiones del código CIE-10. Por ejemplo, Australia presentó su primera edición, la «CIE-10-AM» en 1998; Canadá publicó su versión en el 2000, la «CIE-10-CA». Alemania también tiene su propia extensión, la «CIE-10-GM».

En EE.UU se añadió el anexo con el sistema de clasificación de procedimientos o ICD-10-PCS.

**Condición:** Cada condición de salud puede ser asignada a una categoría y recibir un código de hasta seis caracteres de longitud (en formato de X00.00). Cada una de tales categorías puede incluir un grupo de enfermedades

similares. Los siguientes códigos se utilizan por la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud.

**Diagnóstico Médico:** El diagnóstico o propedéutica clínica es el procedimiento por el cual se identifica una enfermedad, entidad nosológica, síndrome, o cualquier condición de salud-enfermedad (el "estado de salud" también se diagnostica).

En términos de la práctica médica, el diagnóstico es un juicio clínico sobre el estado psicofísico de una persona; representa una manifestación en respuesta a una demanda para determinar tal estado.

Diagnosticar es dar nombre al sufrimiento del paciente; es asignar una "etiqueta".

Tabla 1: *Código de grupo de enfermedades según Clasificación Estadística Internacional de Enfermedades Relacionados con la Salud.*

<b>CÓDIGO</b>	<b>TÍTULO</b>
<a href="#">A00-B99</a>	Ciertas enfermedades infecciosas y parasitarias
<a href="#">C00-D48</a>	Neoplasias
<a href="#">D50-D89</a>	Enfermedades de la sangre y de los órganos hematopoyéticos y otros trastornos que afectan el mecanismo de la inmunidad
<a href="#">E00-E90</a>	Enfermedades endocrinas, nutricionales y metabólicas
<a href="#">F00-F99</a>	Trastornos mentales y del comportamiento
<a href="#">G00-G99</a>	Enfermedades del sistema nervioso
<a href="#">H00-H59</a>	Enfermedades del ojo y sus anexos
<a href="#">H60-H95</a>	Enfermedades del oído y de la apófisis mastoides
<a href="#">I00-I99</a>	Enfermedades del sistema circulatorio
<a href="#">J00-J99</a>	Enfermedades del sistema respiratorio
<a href="#">K00-K93</a>	Enfermedades del aparato digestivo
<a href="#">L00-L99</a>	Enfermedades de la piel y el tejido subcutáneo
<a href="#">M00-M99</a>	Enfermedades del sistema osteomuscular y del tejido conectivo

<a href="#">N00-N99</a>	Enfermedades del aparato genitourinario
<a href="#">O00-O99</a>	Embarazo, parto y puerperio
<a href="#">P00-P96</a>	Ciertas afecciones originadas en el periodo perinatal
<a href="#">Q00-Q99</a>	Malformaciones congénitas, deformidades y anomalías cromosómicas
<a href="#">R00-R99</a>	Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte
<a href="#">S00-T98</a>	Traumatismos, envenenamientos y algunas otras consecuencias de causa externa
<a href="#">V01-Y98</a>	Causas externas de morbilidad y de mortalidad
<a href="#">Z00-Z99</a>	Factores que influyen en el estado de salud y contacto con los servicios de salud
<a href="#">U00-U99</a>	Códigos para situaciones especiales

Fuente: Estructura básica de la CIE 10

### a. Proceso diagnóstico

El diagnóstico clínico requiere tener en cuenta los dos aspectos de la lógica, es decir, el análisis y la síntesis, utilizando diversas herramientas como la anamnesis, la historia clínica, exploración física y exploraciones complementarias.

El diagnóstico médico establece a partir de síntomas, signos y los hallazgos de exploraciones complementarias, qué enfermedad padece una persona. Generalmente una enfermedad no está relacionada de una forma biunívoca con un síntoma, es decir, un síntoma no es exclusivo de una enfermedad. Cada síntoma o hallazgo en una exploración presenta una probabilidad de aparición en cada enfermedad.

El **teorema de Bayes** ayuda al diagnóstico de una enfermedad a partir de los síntomas y otros hallazgos que presenta el paciente si las enfermedades son mutuamente excluyentes, se conoce sus prevalencias y la frecuencia de aparición de cada síntoma en cada enfermedad. Según la prevalencia de cada enfermedad en cada población, un mismo conjunto de síntomas o síndrome

puede producir un diagnóstico diferente en cada población, es decir, cada síndrome puede estar producido por una enfermedad diferente en cada población.

#### **b. Herramientas diagnósticas**

- **Síntomas:** Son las experiencias subjetivas negativas físicas que refiere el paciente, recogidas por el médico en la anamnesis durante la entrevista clínica, con un lenguaje médico, es decir comprensible para todos los médicos. Por ejemplo, los pacientes a la sensación de falta de aire o percepción incómoda y desagradable en la respiración (disnea), lo llaman ahogo, angustia, fatiga, cansancio fácil, etc.
- **Signos:** Son los hallazgos objetivos que detecta el médico observando al paciente, por ejemplo la taquipnea a más de 30 respiraciones por minuto. La semiótica médica o semiología clínica es la parte de la medicina que trata de los signos de las enfermedades desde el punto de vista del diagnóstico y del pronóstico.
- **Exploración física o Semiotécnica:** Consiste en diversas maniobras que realiza el médico sobre el paciente, siendo las principales la inspección, palpación, percusión, oler y auscultación, con las que se obtienen signos clínicos más específicos.

Todos los síntomas referidos en la anamnesis y los signos objetivados en la exploración física son anotados en la historia clínica del paciente.

Generalmente, los signos y síntomas definen un síndrome que puede estar ocasionado por varias enfermedades. El médico debe formular una hipótesis sobre las enfermedades que pueden estar ocasionando el síndrome y para comprobar la certeza de la hipótesis solicita exploraciones complementarias.

Dentro de las exploraciones complementarias confirman o descartan una enfermedad en concreto, antes de iniciar un tratamiento. A veces no ofrecen ningún tipo de información útil, sobre todo cuando se solicitan sin ningún tipo de criterio o no existe un diagnóstico diferencial.

- **Pruebas de laboratorio:** Consiste en el análisis, generalmente bioquímico de diferentes líquidos corporales, siendo el más común la sangre.
- **Técnicas de diagnóstico por imagen:** Como la ecografía, la radiografía simple, la TAC, la RMN, o la PET.
- **Técnicas endoscópicas:**
  - Fibrobroncoscopia
  - Colonoscopia
  - Gastroscopia
  - Colposcopia
  - Toracoscopia
  - Laringoscopia y Biopsia.

### c. Tipos de diagnóstico

#### ▪ **Diagnóstico clínico o individual:**

Es el total emitido a partir del contraste de todos los mencionados y de las condiciones personales del enfermo. Todos estos factores determinan cualitativa y cuantitativamente el cuadro clínico, de manera que éste puede ser diferente aun cuando la entidad morbosa sea la misma. A esto se refiere la máxima «no hay enfermedades, sino enfermos».

#### ▪ **Diagnóstico de certeza:**

Es el diagnóstico confirmado a través de la interpretación y análisis de métodos complementarios.

#### ▪ **Diagnóstico diferencial:**

Conocimiento al que se arriba después de la evaluación crítica comparativa de sus manifestaciones más comunes con las de otras enfermedades.

#### ▪ **Diagnóstico etiológico:**

Determina las causas de la enfermedad; es esencial para el d. total de muchas enfermedades y para el tratamiento.

#### ▪ **Diagnóstico genérico:**

Determinar si el sujeto está o no enfermo. Pueden plantearse problemas de enjuiciamiento, pues hay que contar con una posible simulación y con las neurosis e histerias, que constituyen verdaderas enfermedades.

- **Diagnóstico heroico:**

Cuando el diagnóstico se convierte en una obsesión, en un absurdo, en una especie de imposición mental y profesional, es decir, es un diagnóstico extremo en el sentido de "por sí mismo", e innecesario.

- **Diagnóstico lesional, anatómico o topográfico:**

Es la localización e identificación de las lesiones en los diferentes órganos y tejidos.

- **Diagnóstico nosológico:**

Es la determinación específica de la enfermedad.

- **Diagnóstico patogenético:**

Consigna los mecanismos que producen la enfermedad por la acción de las causas y la reacción orgánica.

- **Diagnóstico presuntivo:**

Es aquel que el profesional considera posible basándose en los datos obtenidos en la anamnesis y el examen físico.

- **Diagnóstico sindrómico y funcional**

Los síndromes son conjuntos de signos y síntomas con un desarrollo común; p. ej., el síndrome icterico (piel amarilla, orinas encendidas, heces decoloradas, etc.). Aunque en algunas ocasiones no se puede avanzar más, permite un d. patogenético parcial, pero que posibilite un tratamiento funcional.

▪ **Diagnóstico sintomático:**

Tiene por objeto identificar la enfermedad mediante los síntomas. Generalmente un síntoma aislado no da una indicación precisa de la enfermedad, puesto que puede ser propio de muchas de ellas.

**d. Problemas del Diagnóstico**

- Diagnóstico tardío
- Diagnóstico ausente
- Diagnóstico genérico
- Diagnóstico inexacto
- Diagnóstico de moda
- Sobrediagnóstico
- Etiquetamiento
- Cascada diagnóstica

**e. Características de una prueba diagnóstica**

Se considera que una prueba diagnóstica es buena cuando ofrece resultados positivos en enfermos y negativos en pacientes sanos, con el menor rango de error posible. Por lo tanto, las condiciones que deben ser exigidas en un test diagnóstico son principalmente tres:

- **Validez:** Es el grado en que un test mide lo que se supone que debe medir, es decir, la frecuencia con la que los resultados obtenidos con este test pueden ser confirmados por otros más complejos y rigurosos. Los parámetros que miden la validez de una prueba diagnóstica son la sensibilidad y la especificidad.

- **Reproductividad:** Es la capacidad de un test de ofrecer los mismos resultados cuando se repite su aplicación en circunstancias similares. La reproductividad viene determinada por la variabilidad biológica del hecho observado, la introducida por el propio observador y la derivada del propio test.
- **Seguridad:** Es la certeza de que un test predecirá la presencia o ausencia de enfermedad en un paciente. La seguridad viene determinada por el valor predictivo de un resultado positivo o negativo, es decir, la probabilidad de que siendo un test positivo el paciente esté realmente enfermo.

#### **f. Validez de una prueba diagnóstica**

El caso más sencillo es el de una prueba dicotómica, la cual clasifica a cada paciente como sano o enfermo, en función de que el resultado de la prueba sea positivo o negativo. De este modo, un resultado positivo se corresponde normalmente con la presencia de la enfermedad estudiada y un resultado negativo con la ausencia de la misma. En general se suele trabajar con una población heterogénea de pacientes, de manera que los datos obtenidos permiten clasificarlos en cuatro grupos, los cuales suelen representarse en una tabla 2X2.

En ella, se enfrenta el resultado de la prueba diagnóstica (en filas) con el estado real de los pacientes (en columnas) o, en su defecto, el resultado de la prueba de referencia o “gold standard” que vayamos a utilizar.

El resultado de la prueba puede ser positivo o negativo, pero estos pueden ser correctos o incorrectos, dando lugar a cuatro tipos de resultados: verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos.

Tabla 2: *Validez de una prueba diagnóstica.*

<b>Resultado de la prueba</b>	<b>Enfermo</b>	<b>Sano</b>
<b>Positivo</b>	Verdaderos positivos (VP)	Falsos positivos (FP)
<b>Negativo</b>	Falsos negativos (FN)	Verdaderos negativos (VN)

**Fuente:** [http://es.wikipedia.org/wiki/Diagn%C3%B3stico\\_m%C3%A9dico#Tipos\\_de\\_diagn.C3.B3stico](http://es.wikipedia.org/wiki/Diagn%C3%B3stico_m%C3%A9dico#Tipos_de_diagn.C3.B3stico)

#### **g. Validez de una prueba diagnóstica**

Lo que determina la validez del test utilizado va a ser el cálculo de los valores de sensibilidad y especificidad.

- **Sensibilidad:** indica la probabilidad de que un paciente esté realmente enfermo siendo el resultado del test positivo. Por tanto, es la capacidad de un test para detectar realmente la presencia de enfermedad. La sensibilidad puede estimarse como la proporción de pacientes enfermos que obtuvieron un resultado positivo en la prueba diagnóstica, es decir, la proporción de verdaderos positivos, o de enfermos diagnosticados, respecto del total de enfermos en la población de estudio. Luego:

$$\textit{Sensibilidad} = \frac{VP}{(VP+FN)}$$

Especificidad: es la probabilidad de que un paciente esté realmente sano tras obtener un resultado del test negativo. Es la capacidad de un test para detectar la ausencia de enfermedad. Así, la especificidad puede estimarse como la proporción de pacientes sanos que obtuvieron un resultado negativo en la prueba diagnóstica, es decir, la proporción de verdaderos negativos, o de sanos reconocidos como tal, respecto del total de sanos en la población. De este modo:

$$\textit{Especificidad} = \frac{VN}{(VN+FP)}$$

Lo ideal es trabajar con pruebas diagnósticas de alta sensibilidad y especificidad, superando el 80% como mínimo en ambos casos. No obstante, esto no siempre es posible. En general, una prueba muy sensible será especialmente adecuada en aquellos casos en los que el no diagnosticar la enfermedad puede resultar fatal para los enfermos, o en enfermedades en las que un falso positivo no produzca serios trastornos psicológicos o económicos para el paciente. En cambio, los test con una alta especificidad son necesarios en enfermedades graves, pero sin tratamiento disponible que las haga curables, cuando exista gran interés por conocer la ausencia de enfermedad o cuando diagnosticar a un paciente de una enfermedad, siendo falso positivo, pueda acarrear graves consecuencias, ya sean físicas, psicológicas o económicas.

#### **h. Eficacia de una prueba diagnóstica**

Tanto la sensibilidad como la especificidad proporcionan información acerca de la probabilidad de obtener un resultado concreto (positivo o negativo) en función de la verdadera condición del paciente con respecto a la enfermedad. Sin embargo, cuando a un paciente se le realiza alguna prueba, el médico carece de información a priori acerca de su verdadero diagnóstico, y más bien la pregunta se plantea en sentido contrario: ante un resultado positivo o negativo en el test, ¿cuál es la probabilidad de que el paciente esté realmente enfermo o sano? Los parámetros que proporcionan esta información (probabilidad post-test) al médico son los denominados valores predictivos.

Los valores predictivos dependerán de la prevalencia de la enfermedad en la población en estudio. Será por lo tanto un valor no extrapolable a poblaciones diferentes. Existen dos tipos de valor predictivo, que se detallan a continuación:

- **Valor predictivo positivo:** indica la probabilidad de que el paciente padezca la enfermedad tras obtener un resultado positivo en el test. Por tanto, el valor predictivo positivo (VPP) puede estimarse como la proporción de verdaderos positivos respecto del total de resultados positivos obtenidos en el test, es decir, el número real de enfermos respecto de todos los resultados que indican presencia de enfermedad.

Así:

$$VPP = \frac{VP}{(VP+FP)}$$

- **Valor predictivo negativo:** es la probabilidad de que el paciente no padezca la enfermedad tras obtener un resultado negativo en el test. De este modo, el valor predictivo negativo (VPN) puede estimarse como la proporción de verdaderos negativos respecto del total de resultados negativos obtenidos en el test, es decir, el número real de pacientes sanos respecto de todos los resultados que indican ausencia de enfermedad. Luego:

$$VPN = \frac{VN}{(VN+FN)}$$

### ***I Elección de una prueba diagnóstica***

Para elegir correctamente entre dos o más pruebas diagnósticas se puede recurrir a parámetros estadísticos. Destaca el conocido como "Característica Operativa del Receptor" o curva ROC. La curva ROC es una representación que compara la sensibilidad de la prueba con el parámetro (1-Especificidad) suponiendo así una medida global e independiente de cualquier punto de corte establecido. El parámetro indicador más utilizado es el "área bajo la curva" (AUC). Se trata de un índice cuyo valor está comprendido entre 0,5 y 1; siendo 1 el valor que determina un diagnóstico perfecto, y 0,5 una prueba sin capacidad.

#### **2.2.2. Definición de términos**

**Neurona:** Una neurona es una unidad de procesamiento de información que es fundamental para el funcionamiento de una red neuronal (Haykin, 2009).

**Aprendizaje supervisado:** Los datos del conjunto de aprendizaje tienen dos tipos de atributos: los datos propiamente dichos y cierta información relativa a la solución del problema (Isasi & Galván, 2004).

**Aprendizaje no supervisado:** Los datos del conjunto de aprendizaje sólo tienen información de los ejemplos, y no hay nada que permita guiar en el proceso de aprendizaje (Isasi & Galván, 2004).

**Error absoluto medio porcentual** (Villada, Muñoz, & García-Quintero, 2016):

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{y'_t - y_t}{y_t} \right|$$

**CAPÍTULO III: HIPÓTESIS Y  
OPERACIONALIZACIÓN DE VARIABLES**

### 3.1. Hipótesis

Si es posible pronosticar las Enfermedades prevalentes mediante Redes Neuronales Artificiales en el Hospital Víctor Ramos Guardia, 2016.

### 3.2. Operacionalización de variables

Las variables de la presente investigación son:

VARIABLES	DIMENSIONES	INDICADORES	ESCALA
<b>Dependiente: Enfermedades Prevalentes</b>	Enfermedades Infecciosas y parasitarias	Número de pacientes con Enfermedades Infecciosas y parasitarias	Razón
	Enfermedades del Sistema respiratorio	Número de pacientes con Enfermedades del Sistema respiratorio	Razón
	Enfermedades Traumatismos, envenenamientos y algunas otras consecuencias de causa externa	Número de pacientes con Enfermedades Traumatismos, envenenamientos y algunas otras consecuencias de causa externa	Razón
	Enfermedades de Factores que influyen en el estado de salud y contacto con los servicios de salud	Número de pacientes con Enfermedades de Factores que influyen en el estado de salud y contacto con los servicios de salud	Razón
	Ciertas enfermedades infecciosas y parasitarias	Número de pacientes con ciertas enfermedades infecciosas y parasitarias	Razón
<b>Independiente:</b>			
t: Tiempo	Tiempo ( Historial de Registro del 2011 al 2016)	Mensual	Razón
<b>VARIABLES INTERVINIENTES</b>	<b>DIMENSIONES</b>	<b>INDICADORES</b>	<b>ESCALA</b>
Edad	Edad cronológica	Años	Razón
Género	Sexo	Masculino / Femenino	Nominal
Procedencia	Lugar donde habita	Distrito	Nominal
Tipo de paciente	Asegurado	Tipo de seguro	Nominal
Motivo de atención	Motivo de ingreso	Causa de la enfermedad	Nominal
Servicio de atención	Servicio médico	Área de referencia médica	Nominal
Destino del paciente	Orden de alta	Lugar de destino	Nominal

## **CAPÍTULO IV: METODOLOGÍA**

## **4.1. Materiales y lugar de ejecución**

### **4.1.1. Lugar**

Hospital Víctor Ramos Guardia de Huaraz.

### **4.1.2. Materiales**

Los materiales utilizados para el presente trabajo de tesis fueron los útiles de escritorio como papel bond, lapiceros, fichas de recolección de datos, USB.

### **4.1.3. Equipos**

Los equipos utilizados en el presente trabajo de tesis fueron una laptop, una impresora, fotocopidora.

### **4.1.4. Recurso humano**

El presente trabajo de tesis cuenta con la asesoría de un docente de la escuela académico profesional de Estadística e Informática y dos investigadores quienes son los autores de la presente tesis.

## **4.2. Métodos**

### **4.2.1. Técnicas de recolección de datos**

Para la presente investigación se recolectarán los registros de atenciones y la base de datos del Hospital Víctor Ramos Guardia bajo la autorización de los responsables del área de Estadística.

#### 4.2.2. Tipo de estudio

- *Según su propósito o finalidad:* Aplicada
- *Según su alcance temporal:* Longitudinal
- *Según su carácter:* cuantitativa
- *Según su nivel de conocimientos que se adquieren o profundidad:* correlacional
- *Según el método utilizado:* Analítica, inductiva y deductiva.

#### 4.2.3. Diseño de la investigación

La presente investigación es de diseño no experimental, longitudinal de tipo serie de tiempo.

#### 4.2.3. Población y muestra

**4.2.3.1. Población:** Está constituida por todos los registros de atenciones del hospital Víctor Ramos Guardia – Huaraz, hasta el 2016.

**4.2.3.2. Muestra:** Para la presente investigación se segmentó los registros entre los años 2011 al 2016, es decir se estudió a todos los registros de atenciones del hospital Víctor Ramos Guardia – Huaraz, periodo 2011 al 2016.

#### 4.2.4. Instrumentos de recopilación de datos

- Técnica:

- Documentaria: Registros de atenciones del HVRG de Huaraz.

- Instrumento:

- Registro de atenciones del HVRG de Huaraz.

#### **4.2.5. Análisis Estadístico e interpretación de la información**

Para el procesamiento y organización de los datos se utilizará el software SPSS v20 y R Project.

El análisis e interpretación estadística de los datos para la presente investigación se realizó por medio de las medidas de tendencia central y dispersión, presentados en gráficos estadísticos y organizados en los cuadros estadísticos. Además, se hizo uso del análisis de redes neuronales artificiales.

# **CAPÍTULO V: ANÁLISIS Y DISCUSIÓN DE RESULTADOS**

## 5.1. Análisis de los datos

### *Resultados descriptivos*

Como se observa en el anexo 2, una de las enfermedades de mayor prevalencia en el Hospital Víctor Ramos Guardia para el periodo 2011 – 2016 son los factores que influyen en el estado de salud y contacto con los servicios de salud (24,3%), seguido por los traumatismos, envenenamientos y algunas otras consecuencias de causa externa (21,0%), las enfermedades del sistema respiratorio (12,9%), los síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte (12,4%) y en quinto lugar ciertas enfermedades infecciosas y parasitarias (8,5%).

Para el análisis de datos de las cinco enfermedades de mayor prevalencia, que representan al 79,1% del total de pacientes del Hospital Víctor Ramos Guardia, 2011 – 2016, se realizó el filtro de los datos utilizando el siguiente código:

```
1
2 DATASET ACTIVATE ConjuntoDatos3.
3 FILTER OFF.
4 USE ALL.
5 SELECT IF (grudiagnost = 1 | grudiagnost = 10 | grudiagnost = 19 | grudiagnost = 21 | grudiagnost =
6 18).
7 EXECUTE.
8
```

Dicho filtro redujo de un total de 186293 pacientes a 147321 pacientes de las cinco enfermedades antes mencionadas.

Del anexo 3, se observa que el 56,0% de los pacientes son de sexo femenino, mientras que el 44,0% son de sexo masculino. El 80,1% de los pacientes de sexo femenino presentan enfermedades de los factores que influyen en el estado de salud y contacto con los servicios de salud, el 65,2% de los pacientes de sexo masculino

presentan enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa, el 51,8% de los pacientes de sexo masculino presentan enfermedades del sistema respiratorio, el 55,4% de los pacientes de sexo femenino presentan enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte, y el 52,1% de los pacientes de sexo femenino presentan ciertas enfermedades infecciosas y parasitarias.

Del anexo 4, se observa que el 20,7% de los pacientes son tienen de 4 a menos años de edad, mientras que el 12,9% tienen entre 20 a 24 años de edad. El 22,0% de los pacientes que tienen entre 20 a 24 años de edad presentan enfermedades de los factores que influyen en el estado de salud y contacto con los servicios de salud, el 12,2% de los pacientes que tienen entre 20 a 24 años de edad presentan enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa, el 49,9% de los pacientes que tienen entre 0 a 4 años de edad presentan enfermedades del sistema respiratorio, el 24,3% de los pacientes que tienen entre 0 a 4 años de edad presentan enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte, y el 32,2% de los pacientes que tienen entre 0 a 4 años de edad presentan ciertas enfermedades infecciosas y parasitarias. Del anexo 5, se observa que el 99,1% de los pacientes no cuentan con ningún tipo de seguro, mientras que el 0,5% cuentan con el SIS. El 99,0% de los pacientes que no tienen seguro de salud presentan enfermedades de los factores que influyen en el estado de salud y contacto con los servicios de salud, el 99,3% de los pacientes que no tienen seguro de salud presentan enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de

causa externa, el 99,1% de los pacientes que no tienen seguro de salud presentan enfermedades del sistema respiratorio, el 99,0% de los pacientes que no tienen seguro de salud presentan enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte, y el 99,3% de los pacientes que no tienen seguro de salud presentan ciertas enfermedades infecciosas y parasitarias.

Del anexo 6, se observa que el 70,9% de los pacientes ingresaron al hospital por causa de enfermedad común, mientras que el 24,6% ingresaron al hospital por causa de enfermedad súbita. El 86,6% de los pacientes que ingresaron al hospital por causa de enfermedad común presentan enfermedades de los factores que influyen en el estado de salud y contacto con los servicios de salud, el 79,2% de los pacientes que ingresaron al hospital por causa de enfermedad súbita presentan enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa, el 97,7% de los pacientes que ingresaron al hospital por causa de enfermedad común presentan enfermedades del sistema respiratorio, el 96,5% de los pacientes que ingresaron al hospital por causa de enfermedad común presentan enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte, y el 98,0% de los pacientes que ingresaron al hospital por causa de enfermedad común presentan ciertas enfermedades infecciosas y parasitarias. Del anexo 7, se observa que el 28,3% de los pacientes ingresaron al hospital por el servicio de medicina general, mientras que el 27,7% ingresaron al hospital por el servicio de cirugía. El 59,1% de los pacientes que ingresaron al hospital por el servicio de obstetricia presentan enfermedades de los factores que influyen en el

estado de salud y contacto con los servicios de salud, el 85,9% de los pacientes que ingresaron al hospital por el servicio de cirugía presentan enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa, el 64,4% de los pacientes que ingresaron al hospital por el servicio de pediatría general presentan enfermedades del sistema respiratorio, el 60,0% de los pacientes que ingresaron al hospital por el servicio de medicina general presentan enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte, y el 58,2% de los pacientes que ingresaron al hospital por el servicio de medicina general presentan ciertas enfermedades infecciosas y parasitarias.

Del anexo 8, se observa que el 71,0% de los pacientes fueron destinados a su domicilio, mientras que el 15,2% fueron destinados a hospitalización. El 40,5% de los pacientes que fueron destinados a su domicilio presentan enfermedades de los factores que influyen en el estado de salud y contacto con los servicios de salud, el 80,0% de los pacientes que fueron destinados a su domicilio presentan enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa, el 90,6% de los pacientes que fueron destinados a su domicilio presentan enfermedades del sistema respiratorio, el 82,7% de los pacientes que fueron destinados a su domicilio presentan enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte, y el 89,6% de los pacientes que fueron destinados a su domicilio presentan ciertas enfermedades infecciosas y parasitarias.

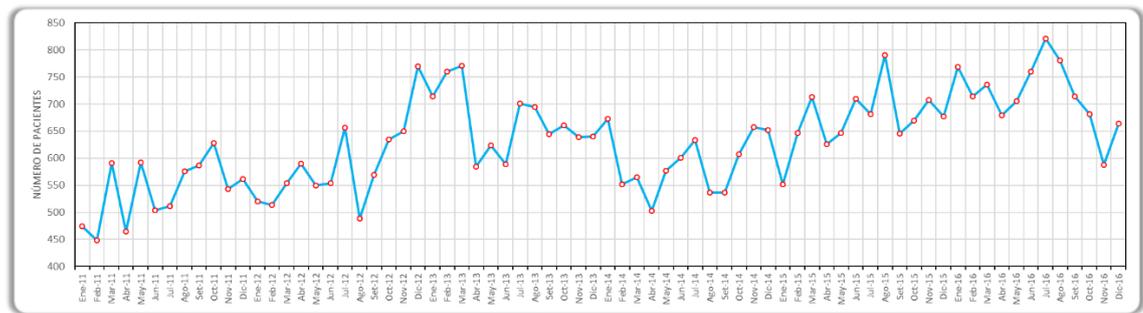
De la tabla 3 y figura 1, se observa que el número de pacientes con enfermedades de factores que influyen en el estado de salud y contacto con los servicios de salud del Hospital Víctor Ramos Guardia, 2011 – 2016, presenta una tendencia creciente, sin apreciarse formas de estacionalidad, con variaciones oscilantes a lo largo del tiempo. Así mismo, el pico más alto de pacientes se encuentra en julio del 2016, con un total de 821 casos, mientras que el pico más bajo sucedió el febrero del 2011 con un total de 448 casos.

Tabla 3

*Número de pacientes con enfermedades de factores que influyen en el estado de salud y contacto con los servicios de salud del Hospital Víctor Ramos Guardia, 2011 – 2016*

<b>Mes</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>
Enero	475	520	714	673	552	769
Febrero	448	514	760	552	647	714
Marzo	591	554	771	565	713	736
Abril	465	590	585	503	626	679
Mayo	592	550	624	577	647	705
Junio	504	554	589	601	710	760
Julio	511	656	701	634	681	821
Agosto	576	489	695	537	790	781
Setiembre	587	569	644	537	645	714
Octubre	628	635	661	607	670	682
Noviembre	543	650	639	658	708	588
Diciembre	562	770	640	652	677	664

Fuente: Elaboración propia



Fuente: Elaboración propia

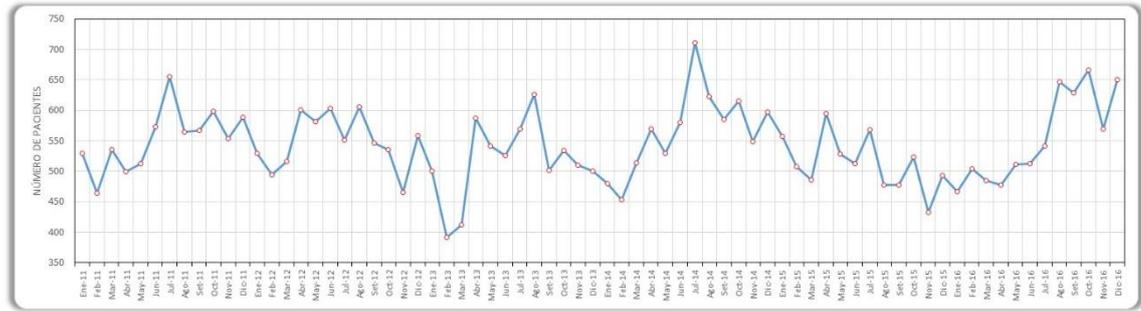
*Figura 1.* Número de pacientes con enfermedades de factores que influyen en el estado de salud y contacto con los servicios de salud del Hospital Víctor Ramos Guardia, 2011 – 2016.

Tabla 4

*Número de pacientes con enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa del Hospital Víctor Ramos Guardia, 2011 – 2016*

Mes	2011	2012	2013	2014	2015	2016
Enero	529	530	500	480	557	466
Febrero	464	494	391	453	508	504
Marzo	535	516	412	514	486	485
Abril	499	601	587	570	595	478
Mayo	512	582	541	530	528	511
Junio	573	603	526	580	513	512
Julio	655	551	570	711	568	541
Agosto	565	606	626	623	477	647
Setiembre	567	547	502	585	478	629
Octubre	598	535	534	616	524	666
Noviembre	554	465	510	549	433	570
Diciembre	589	558	500	597	493	650

Fuente: Elaboración propia



Fuente: Elaboración propia

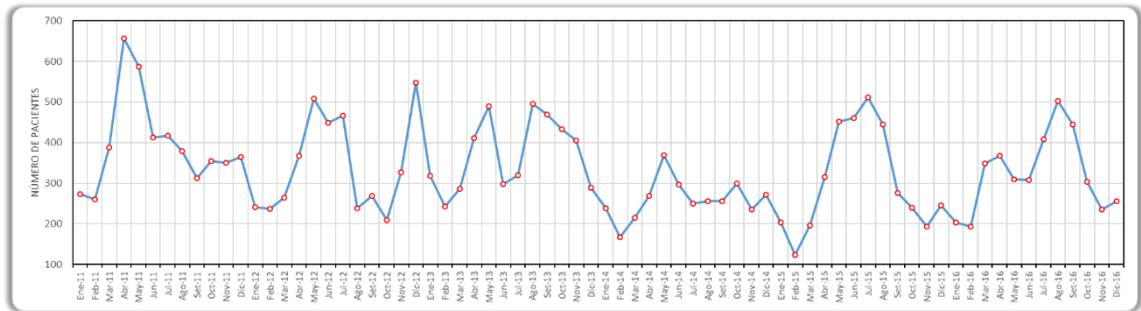
*Figura 2.* Número de pacientes con enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa del Hospital Víctor Ramos Guardia, 2011 – 2016.

Tabla 5

*Número de pacientes con enfermedades del sistema respiratorio del Hospital Víctor Ramos Guardia, 2011 – 2016*

Mes	2011	2012	2013	2014	2015	2016
Enero	273	241	318	238	204	203
Febrero	260	237	243	167	123	194
Marzo	388	264	286	215	196	349
Abril	657	368	411	269	316	368
Mayo	587	508	489	369	452	310
Junio	413	449	298	297	461	308
Julio	417	467	320	250	511	408
Agosto	379	238	495	256	444	502
Setiembre	312	269	469	256	276	444
Octubre	355	210	433	300	240	304
Noviembre	350	327	405	235	194	235
Diciembre	365	548	289	272	245	256

Fuente: Elaboración propia



Fuente: Elaboración propia

*Figura 3.* Número de pacientes con enfermedades del sistema respiratorio del Hospital Víctor Ramos Guardia, 2011 – 2016.

De la tabla 4 y figura 2, se observa que el número de pacientes con enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa del Hospital Víctor Ramos Guardia, 2011 – 2016, presenta una tendencia ligeramente creciente, sin apreciarse formas de estacionalidad, con variaciones oscilantes a lo largo del tiempo. Así mismo, el pico más alto de pacientes se encuentra en julio del 2014, con un total de 711 casos, mientras que el pico más bajo sucedió el febrero del 2013 con un total de 391 casos.

De la tabla 5 y figura 3, se observa que el número de pacientes con enfermedades del sistema respiratorio del Hospital Víctor Ramos Guardia, 2011 – 2016, no evidencia presencia de una tendencia, sin apreciarse formas de estacionalidad, con variaciones oscilantes a lo largo del tiempo. Así mismo, el pico más alto de pacientes se encuentra en abril del 2011, con un total de 657 casos, mientras que el pico más bajo sucedió el febrero del 2015 con un total de 123 casos.

De la tabla 6 y figura 4, se observa que el número de pacientes con enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte del Hospital Víctor Ramos Guardia, 2011 – 2016, evidencia presencia de una tendencia creciente, sin apreciarse formas de estacionalidad, con variaciones suavizadas oscilantes a lo largo del tiempo. Así mismo, el pico más alto de pacientes se encuentra en agosto del 2016, con un total de 516 casos, mientras que el pico más bajo sucedió el febrero del 2011 con un total de 191 casos.

Tabla 6

*Número de pacientes con enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte del Hospital Víctor Ramos Guardia, 2011 – 2016*

<b>Mes</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>
Enero	267	298	294	283	320	393
Febrero	191	246	249	229	267	282
Marzo	254	265	296	257	258	339
Abril	319	309	375	258	363	391
Mayo	337	284	354	291	403	433
Junio	277	297	333	283	370	384
Julio	329	352	304	315	371	379
Agosto	287	340	353	299	436	516
Setiembre	309	312	295	337	399	508
Octubre	266	301	263	334	319	393
Noviembre	252	327	292	317	327	374
Diciembre	274	303	270	279	299	394

Fuente: Elaboración propia



Fuente: Elaboración propia

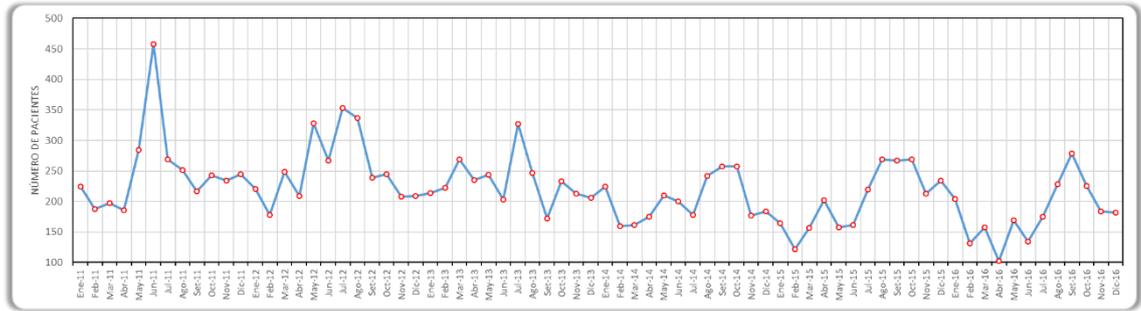
*Figura 4.* Número de pacientes con enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte del Hospital Víctor Ramos Guardia, 2011 – 2016.

Tabla 7

*Número de pacientes con ciertas enfermedades infecciosas y parasitarias del Hospital Víctor Ramos Guardia, 2011 – 2016*

<b>Mes</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>
Enero	224	221	214	224	164	204
Febrero	188	178	222	160	122	131
Marzo	197	249	269	161	157	158
Abril	186	209	235	175	202	102
Mayo	284	328	244	210	158	169
Junio	458	267	203	200	161	134
Julio	269	353	327	178	220	175
Agosto	252	337	247	242	269	228
Setiembre	217	239	172	257	267	279
Octubre	243	245	233	257	269	225
Noviembre	234	208	213	177	213	184
Diciembre	245	209	206	184	234	182

Fuente: Elaboración propia



Fuente: Elaboración propia

*Figura 5.* Número de pacientes con ciertas enfermedades infecciosas y parasitarias del Hospital Víctor Ramos Guardia, 2011 – 2016.

De la tabla 7 y figura 5, se observa que el número de pacientes con ciertas enfermedades infecciosas y parasitarias del Hospital Víctor Ramos Guardia, 2011 – 2016, evidencia presencia de una ligera tendencia decreciente, sin apreciarse formas de estacionalidad, con variaciones oscilantes a lo largo del tiempo. Así mismo, el pico más alto de pacientes se encuentra en junio del 2011, con un total de 458 casos, mientras que el pico más bajo sucedió el abril del 2016 con un total de 102 casos.

***Análisis de Redes Neuronales Artificiales***

Para el análisis de redes neuronales artificiales, se procedió a realizar los siguientes pasos:

Primero: Establecemos nuestro directorio y cargamos los datos en el entorno R:

```
library(readxl)

Serie<-read_excel("Serie.xlsx", col_types = c("date", "date",
"numeric"))

attach(Serie)
```

Segundo: Uno de los procedimientos más importantes cuando se forma una red neuronal es la normalización de datos. Esto implica ajustar los datos a una escala común para comparar con precisión los valores predichos y reales.

```
normalize <- function(x) {
  return ((x - min(x)) / (max(x) - min(x)))
}
```

Tercero: Basamos nuestros datos de entrenamiento en el 80% de las observaciones.

Los datos de la prueba (validación) se basan en el 20% restante de las

```
trainset <- maxmindf[1:48, ]
testset <- maxmindf[49:67, ]
```

observaciones.

Cuarto: El entrenamiento del modelo de red neuronal artificial se realizó mediante la función “neuralnet”

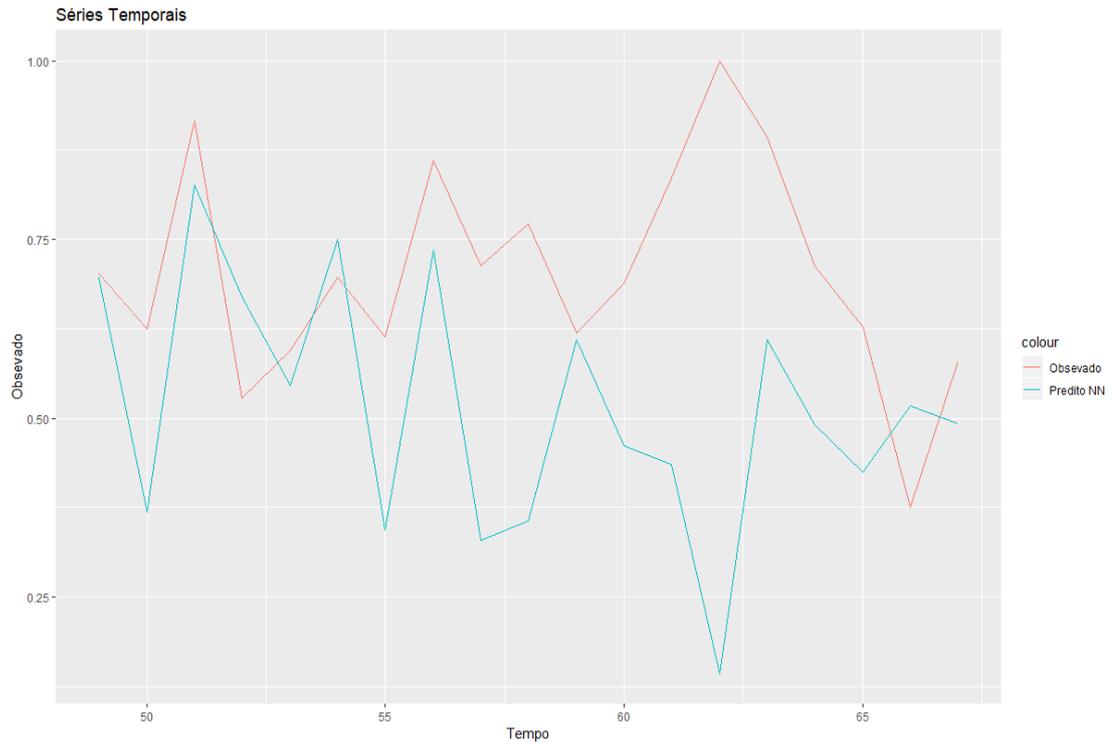
```
set.seed(12345)
library("neuralnet")
maxit<-as.integer(1000000)
nn<-
neuralnet(t5~t1+t2+t3+t4, data=Serie.Entrena, hidden=c(8, 6), threshold
=0.01, stepmax=maxit)
plot(nn)
pronostico<-compute(nn, Serie.Valida[,1:4])
pronostico.nn<-pronostico$net.result
```

Finalmente, los resultados para las cinco enfermedades más prevalentes de los pacientes del Hospital Víctor Ramos Guardia, 2011 – 2016, fueron obtenidos con los códigos anteriores.

**Pacientes con enfermedades de factores que influyen en el estado de salud y contacto con los servicios de salud del Hospital Víctor Ramos Guardia, 2011 – 2016**

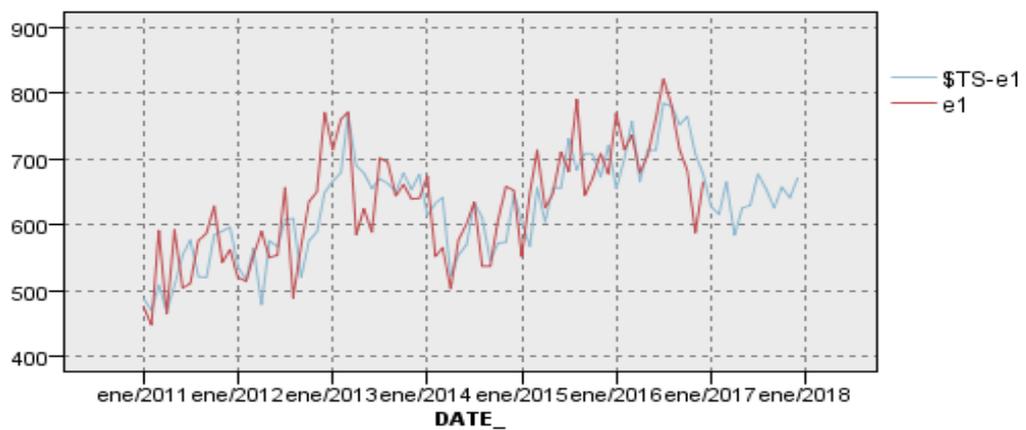
El modelo de red neuronal artificial para esta variable consta de una red con dos capas ocultas, compuesta por ocho neuronas en la primera capa y seis neuronas en la segunda capa, con un umbral de 0.01, lo que significa que, si el cambio en el error durante una iteración es inferior al 1%, el modelo no llevará a cabo ninguna otra optimización. Además, se incluyó cuatro rezagos de la serie temporal analizada.





Fuente: Elaboración propia

*Figura 7.* Muestra de validación de la red neuronal para el número de pacientes con enfermedades de factores que influyen en el estado de salud y contacto con los servicios de salud del Hospital Víctor Ramos Guardia, 2011 – 2016.



Fuente: Elaboración propia

*Figura 8.* Serie real y pronosticada con la red neuronal para el número de pacientes con enfermedades de factores que influyen en el estado de salud y contacto con los servicios de salud del Hospital Víctor Ramos Guardia, 2011 – 2016.

Tabla 8

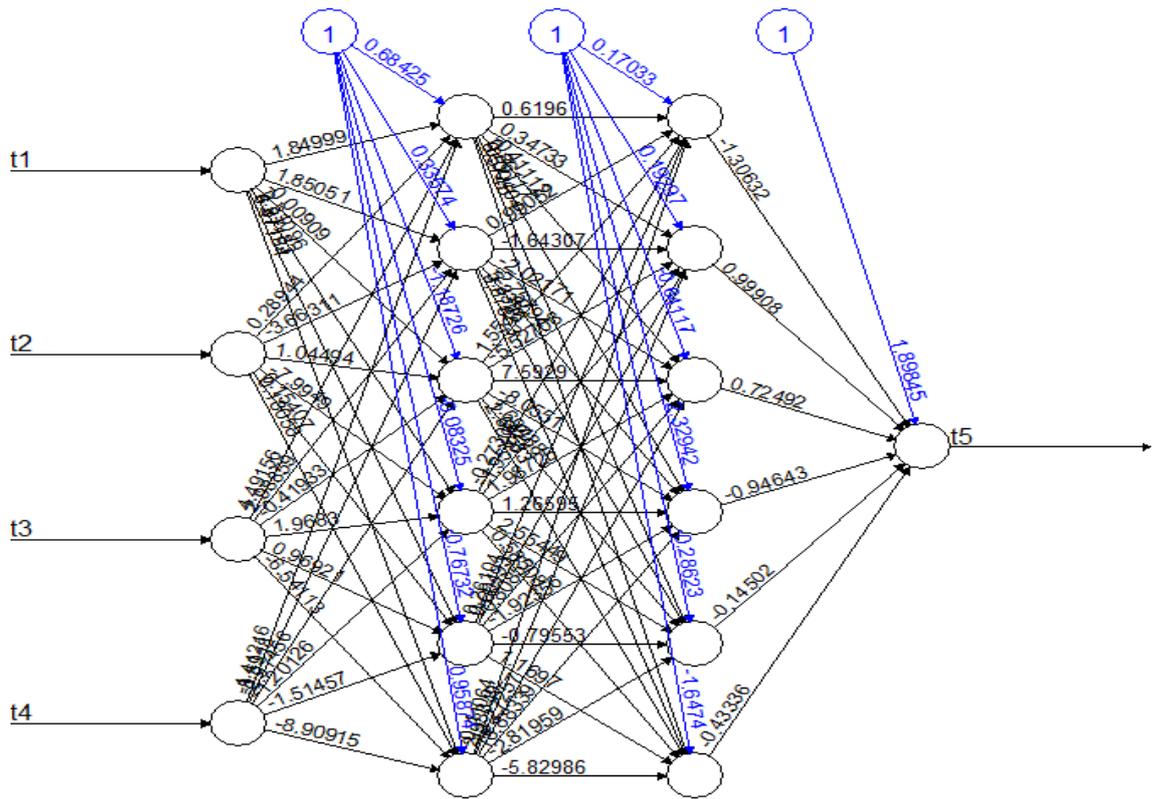
*Pronóstico del número de pacientes con enfermedades de factores que influyen en el estado de salud y contacto con los servicios de salud del Hospital Víctor Ramos Guardia, periodo 2017*

<b>Tiempo</b>	<b>Pronóstico</b>
Ene-17	626.961
Feb-17	615.627
Mar-17	664.794
Abr-17	584.461
May-17	625.627
Jun-17	629.461
Jul-17	677.127
Ago-17	654.461
Set-17	625.794
Oct-17	656.961
Nov-17	640.794
Dic-17	670.627

Fuente: Elaboración propia

**Pacientes con enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa del Hospital Víctor Ramos Guardia, 2011 – 2016**

El modelo de red neuronal artificial para esta variable consta de una red con dos capas ocultas, compuesta por seis neuronas en la primera capa y seis neuronas en la segunda capa, con un umbral de 0.01, lo que significa que, si el cambio en el error durante una iteración es inferior al 1%, el modelo no llevará a cabo ninguna otra optimización. Además, se incluyó cuatro rezagos de la serie temporal analizada.

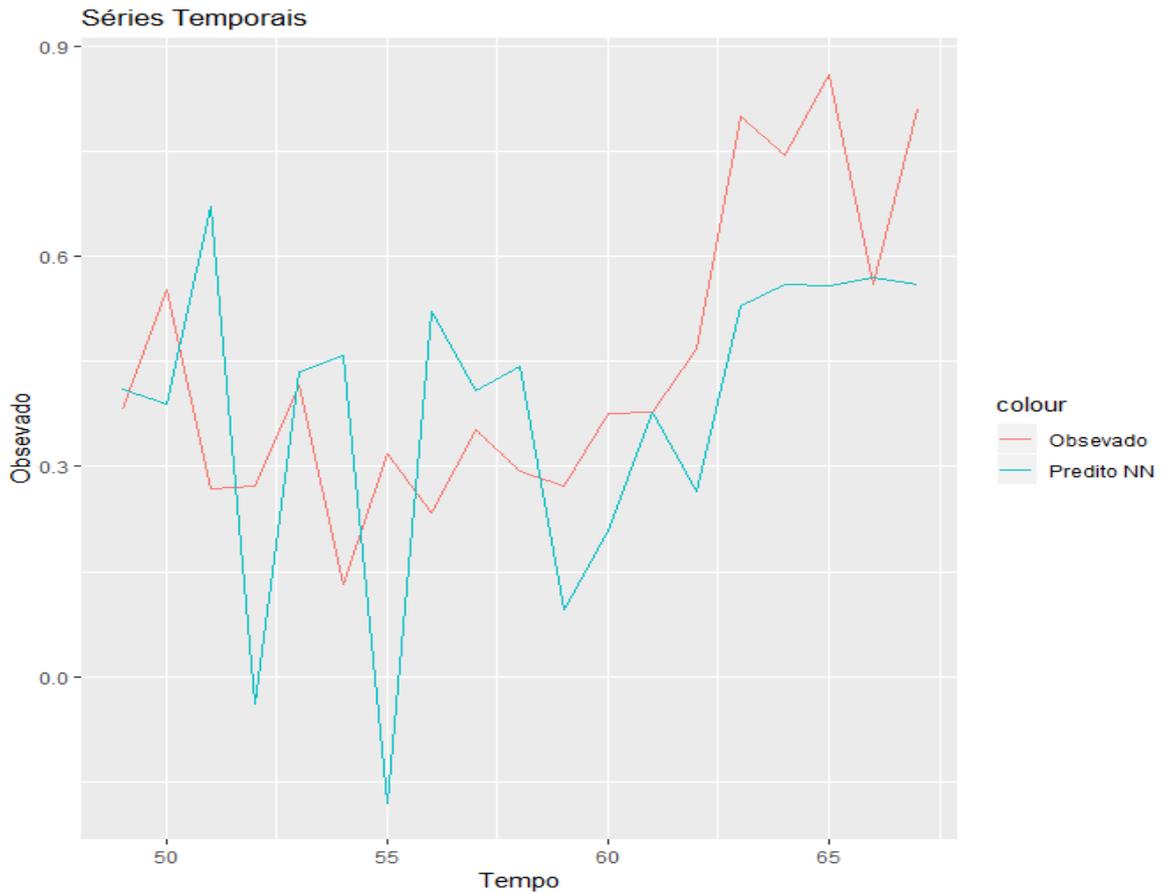


Error: 0.370842 Steps: 347

Fuente: Elaboración propia

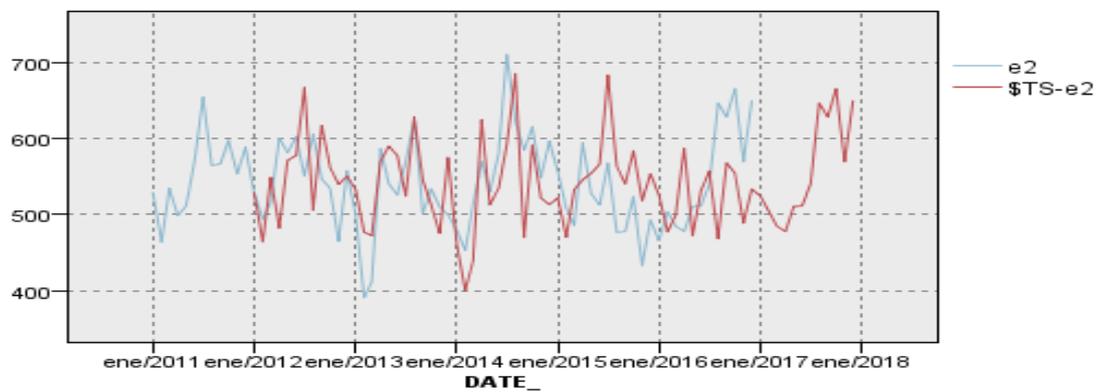
Figura 9. Modelo de red neuronal para el número de pacientes con enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa del Hospital Víctor Ramos Guardia, 2011 – 2016.

El error de entrenamiento es de 0.370842 en 347 iteraciones, con un umbral del 0.00900121 (0.900121%) para el ajuste de las ponderaciones.



Fuente: Elaboración propia

Figura 10. Muestra de validación de la red neuronal para el número de pacientes con enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa del Hospital Víctor Ramos Guardia, 2011 – 2016.



Fuente: Elaboración propia

*Figura 11.* Serie real y pronosticada con la red neuronal para el número de pacientes con enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa del Hospital Víctor Ramos Guardia, 2011 – 2016.

Tabla 9

*Pronóstico del número de pacientes con enfermedades de traumatismos, envenenamientos y algunas otras consecuencias de causa externa del Hospital Víctor Ramos Guardia, periodo 2017*

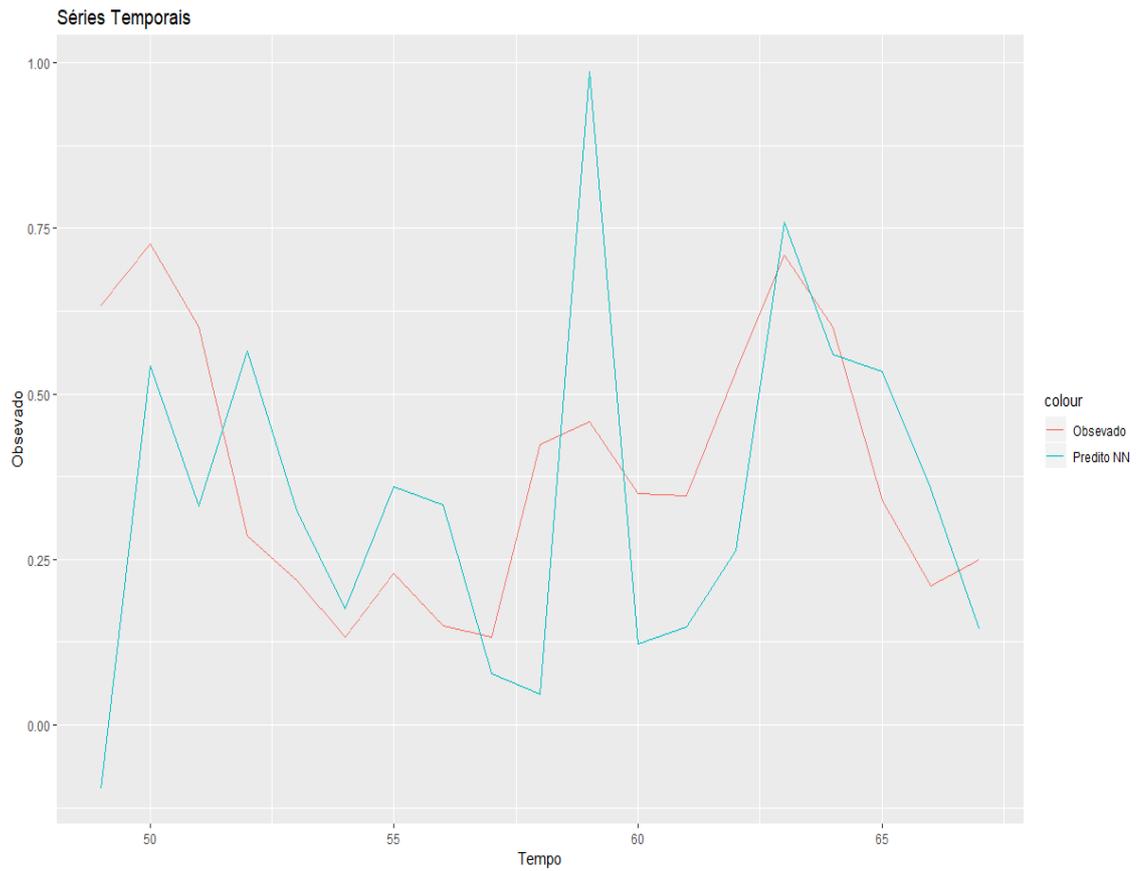
<b>Tiempo</b>	<b>Pronóstico</b>
Ene-17	524.751
Feb-17	504.000
Mar-17	485.000
Abr-17	478.000
May-17	511.000
Jun-17	512.000
Jul-17	541.000
Ago-17	647.000
Set-17	629.000
Oct-17	666.000
Nov-17	570.000
Dic-17	650.000

Fuente: Elaboración propia

### **Número de pacientes con enfermedades del sistema respiratorio del Hospital Víctor Ramos Guardia, 2011 – 2016**

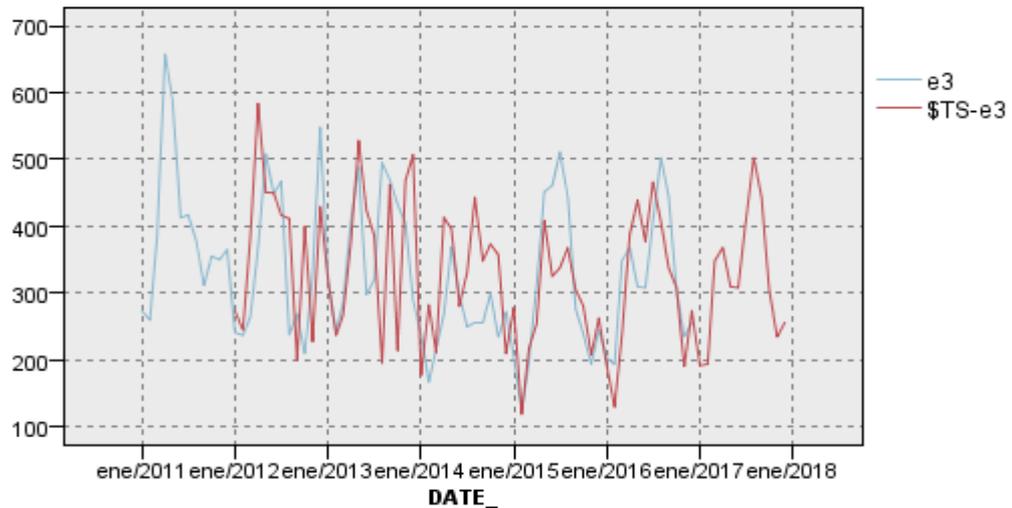
El modelo de red neuronal artificial para esta variable consta de una red con dos capas ocultas, compuesta por siete neuronas en la primera capa y siete neuronas en la segunda capa, con un umbral de 0.01, lo que significa que, si el cambio en el error durante una iteración es inferior al 1%, el modelo no llevará a cabo ninguna otra optimización. Además, se incluyó cuatro rezagos de la serie temporal analizada.





Fuente: Elaboración propia

*Figura 13.* Muestra de validación de la red neuronal para el número de pacientes con enfermedades del sistema respiratorio del Hospital Víctor Ramos Guardia, 2011 – 2016.



Fuente: Elaboración propia

*Figura 14.* Serie real y pronosticada con la red neuronal para el número de pacientes con enfermedades del sistema respiratorio del Hospital Víctor Ramos Guardia, 2011 – 2016.

Tabla 10

*Pronóstico del número de pacientes con enfermedades del sistema respiratorio del Hospital Víctor Ramos Guardia, periodo 2017*

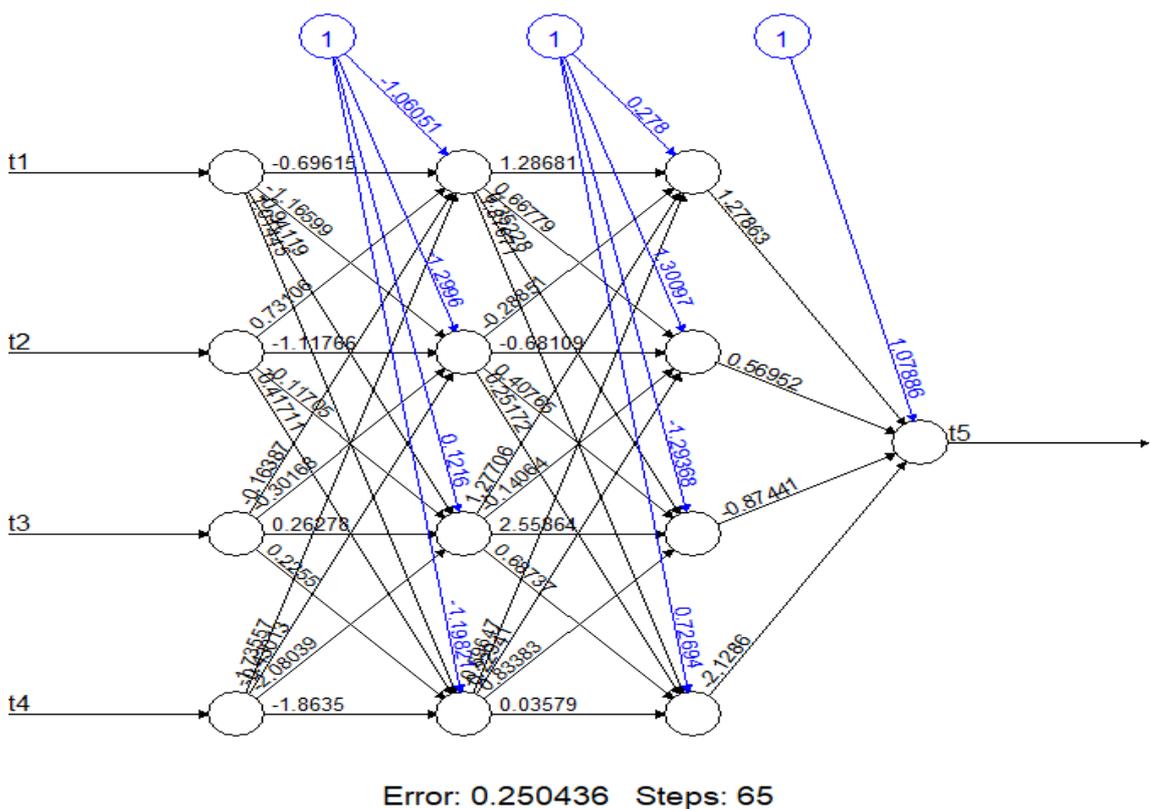
<b>Tiempo</b>	<b>Pronóstico</b>
Ene-17	191.732
Feb-17	194.000
Mar-17	349.000
Abr-17	368.000
May-17	310.000
Jun-17	308.000
Jul-17	408.000
Ago-17	502.000
Set-17	444.000
Oct-17	304.000
Nov-17	235.000
Dic-17	256.000

Fuente: Elaboración propia

**Número de pacientes con enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte del Hospital**

**Víctor Ramos Guardia, 2011 – 2016**

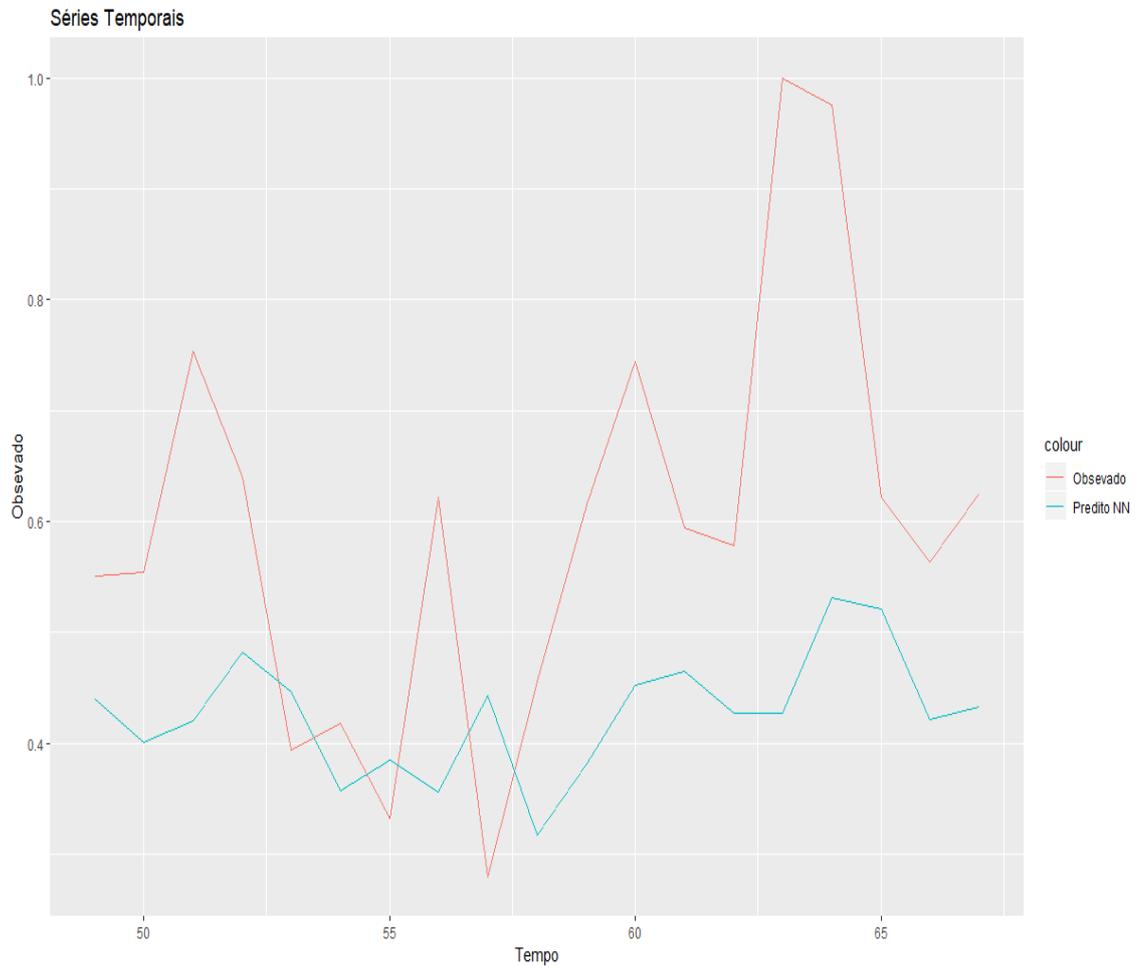
El modelo de red neuronal artificial para esta variable consta de una red con dos capas ocultas, compuesta por cuatro neuronas en la primera capa y cuatro neuronas en la segunda capa, con un umbral de 0.01, lo que significa que, si el cambio en el error durante una iteración es inferior al 1%, el modelo no llevará a cabo ninguna otra optimización. Además, se incluyó cuatro rezagos de la serie temporal analizada.



Fuente: Elaboración propia

*Figura 15.* Modelo de red neuronal para el número de pacientes con enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte del Hospital Víctor Ramos Guardia, 2011 – 2016.

El error de entrenamiento es de 0.250436 en 65 iteraciones, con un umbral del 0.005679315 (0.5679315%) para el ajuste de las ponderaciones



Fuente: Elaboración propia

*Figura 16.* Muestra de validación de la red neuronal para el número de pacientes con enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte del Hospital Víctor Ramos Guardia, 2011 – 2016.



Fuente: Elaboración propia

*Figura 17.* Serie real y pronosticada con la red neuronal para el número de pacientes con enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte del Hospital Víctor Ramos Guardia, 2011 – 2016.

Tabla 11

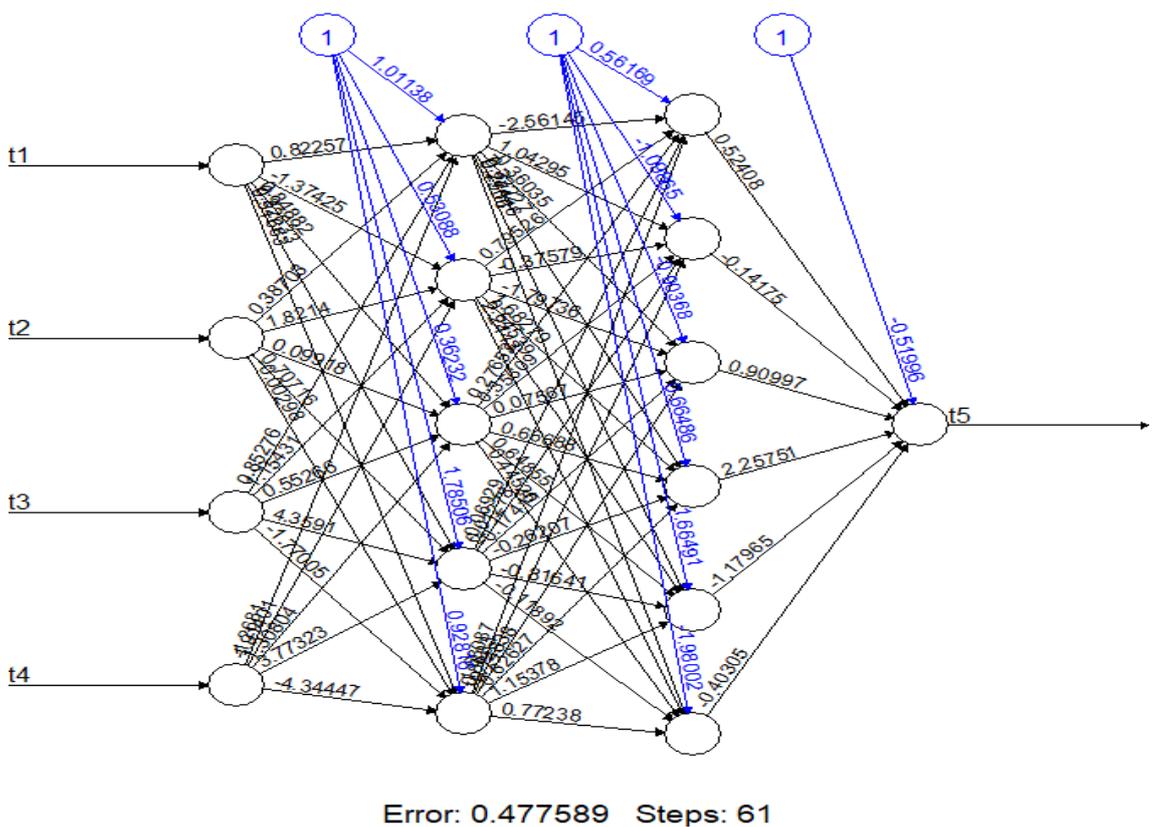
*Pronóstico del número de pacientes con enfermedades de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte del Hospital Víctor Ramos Guardia, periodo 2017*

<b>Tiempo</b>	<b>Pronóstico</b>
Ene-17	443.668
Feb-17	362.421
Mar-17	385.564
Abr-17	464.752
May-17	505.726
Jun-17	464.934
Jul-17	463.012
Ago-17	563.077
Set-17	540.200
Oct-17	443.155
Nov-17	438.005
Dic-17	433.382

Fuente: Elaboración propia

## Número de pacientes con enfermedades de ciertas enfermedades infecciosas y parasitarias del Hospital Víctor Ramos Guardia, 2011 – 2016

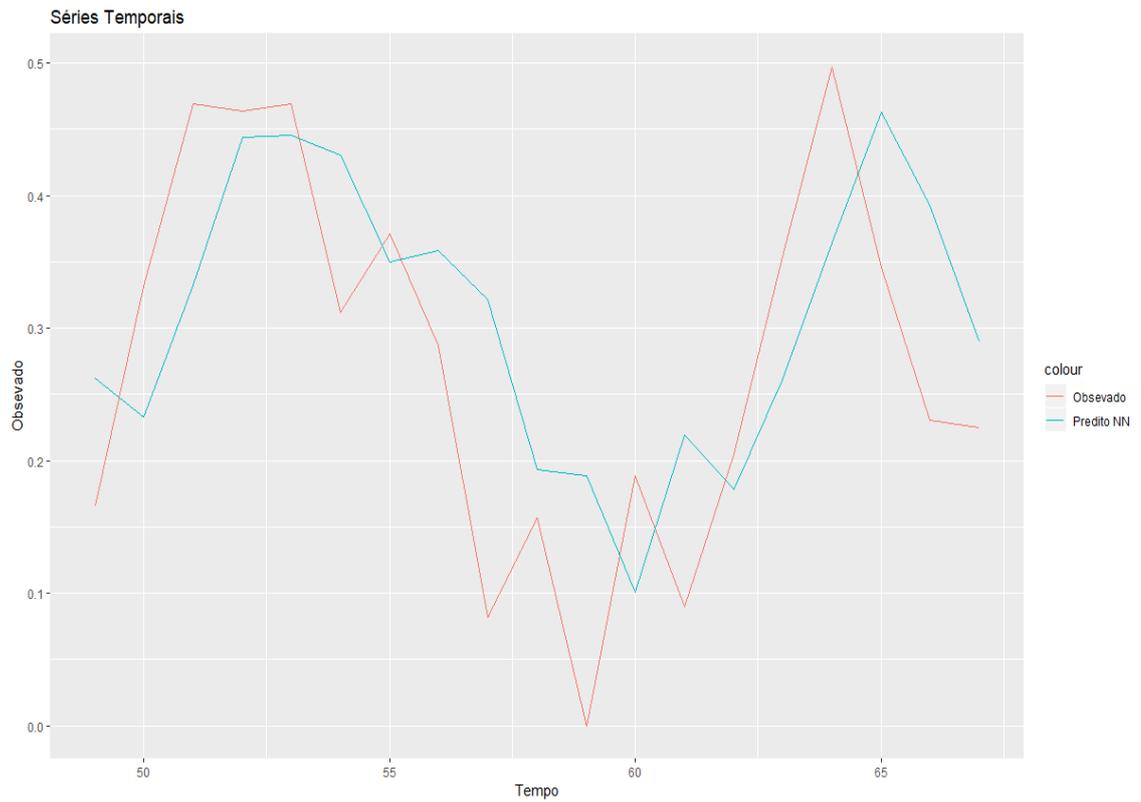
El modelo de red neuronal artificial para esta variable consta de una red con dos capas ocultas, compuesta por cinco neuronas en la primera capa y seis neuronas en la segunda capa, con un umbral de 0.01, lo que significa que, si el cambio en el error durante una iteración es inferior al 1%, el modelo no llevará a cabo ninguna otra optimización. Además, se incluyó cuatro rezagos de la serie temporal analizada.



Fuente: Elaboración propia

Figura 18. Modelo de red neuronal para el número de pacientes con enfermedades de ciertas enfermedades infecciosas y parasitarias del Hospital Víctor Ramos Guardia, 2011 – 2016.

El error de entrenamiento es de 0.477589 en 61 iteraciones, con un umbral del 0.009409171 (0.9409171%) para el ajuste de las ponderaciones.



Fuente: Elaboración propia

*Figura 19.* Muestra de validación de la red neuronal para el número de pacientes con enfermedades de ciertas enfermedades infecciosas y parasitarias del Hospital Víctor Ramos Guardia, 2011 – 2016.



Fuente: Elaboración propia

*Figura 20.* Serie real y pronosticada con la red neuronal para el número de pacientes con enfermedades de ciertas enfermedades infecciosas y parasitarias del Hospital Víctor Ramos Guardia, 2011 – 2016.

Tabla 12

*Pronóstico del número de pacientes con enfermedades de ciertas enfermedades infecciosas y parasitarias del Hospital Víctor Ramos Guardia, periodo 2017*

<b>Tiempo</b>	<b>Pronóstico</b>
Ene-17	158.310
Feb-17	116.643
Mar-17	148.310
Abr-17	134.643
May-17	181.977
Jun-17	186.977
Jul-17	203.477
Ago-17	212.310
Set-17	188.310
Oct-17	195.143
Nov-17	154.643
Dic-17	159.810

Fuente: Elaboración propia

## 5.2. Discusión de resultados

Las enfermedades prevalentes en el Hospital Víctor Ramos Guardia de Huaraz, son caracterizados por enfermedades categorizados como factores que influyen en el estado de salud y contacto con los servicios de salud (E1: 24,3%), traumatismos, envenenamientos y algunas otras consecuencias de causa externa (E2: 21,0%), enfermedades del sistema respiratorio (E3: 12,9%), síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte (E4: 12,4%) y ciertas enfermedades infecciosas y parasitarias (E5: 8,5%). Éstas cinco categorías de enfermedades abarcan el 79,1% del total de enfermedades que fueron atendidos en el hospital, que en su mayoría dichos pacientes son de sexo femenino (56,0%), habitualmente menores de 5 años de edad (20,7%), casi en su totalidad siendo pacientes sin seguro (99,1%), que generalmente fueron atendidos por motivos de enfermedad común (70,9%), fueron atendidos en el servicio de medicina general (28,3%) y que su destino luego de la atención fue derivado a su domicilio (71,0%). Estos resultados concuerdan con los encontrados por Durand & Gustavo (2003) en la que concluyen que las enfermedades bucales son la primera causa de enfermedad tanto a nivel general como en los diferentes grupos estudiados; seguida de la parasitosis y la desnutrición.

Por otro lado, el análisis del algoritmo de una red neuronal artificial está enmarcado en el aprendizaje supervisado, con modelo de perceptrón multicapa con función de activación e integración la función logística con el algoritmo de retropropagación para la corrección de los errores estimados, tal como fue usado por Trujilano (2005). Así mismo, fue posible el modelamiento mediante redes neuronales artificiales a

los datos no lineales, obteniéndose pronósticos con medidas del error inferiores a la unidad, garantizando de esa manera un buen pronóstico, tal como lo indica Delgado (1999), creando sistemas expertos que enriquecen el análisis de las enfermedades brindando nuevas perspectivas a los especialistas.

Así mismo, la red neuronal artificial modelada depende directamente de la cantidad de capas y neuronas que contienen cada capa, teniendo en cuenta que el modelado de la red neuronal está compuesto por tres capas, donde la capa de entrada para el presente estudio estuvo compuesta por cuatro capas (que representan las cuatro series de tiempo rezagadas hasta cuatro periodos de estudio). La cantidad de neuronas está directamente relacionado a la convergencia del umbral de 1% para el inicio de corrección de los errores de estimación de las series E1, E2, E3, E4 y E5, que representan a las enfermedades prevalentes antes encontradas. Así, por ejemplo, en el estudio de Asnate (2014) se modeló una red neuronal artificial con tres capas, la capa de entrada está compuesta con 29 neuronas, la capa oculta está compuesta por 15 neuronas y la salida compuesta por dos neuronas. Cabe indicar que los modelos de redes neuronales artificiales encontradas no son las únicas en su arquitectura, ya que con otras arquitecturas puedan variar los indicadores de error de estimación, así como el pronóstico de las series de tiempo de las enfermedades prevalentes del Hospital Víctor Ramos Guardia, tal como lo manifiesta Collantes (2001) en la que resalta que una de las limitaciones de la red neuronal artificial para alcanzar un uso generalizado es el establecimiento de la arquitectura de la red.

Finalmente, tal como lo señala Bernal (1999), la identificación de un tipo específico de enfermedad es muchas más precisa y sencilla en la actualidad, puesto que muchos de los sistemas de diagnósticos que se aplican hoy en día, poseen una gran cantidad de conocimiento (plasmado generalmente como bases de datos) realizado con las redes neuronales artificiales, además del hecho que este conocimiento es útil para el reconocimiento de enfermedades que presentes características comunes y de ésta manera ampliar la base de conocimiento presente, lo cual coincide con los resultados del presente estudio, en cuanto al uso de herramientas de análisis actuales y aprovechando las características de su uso y aplicación como son las redes neuronales artificiales, con convergencia a dos capas ocultas y errores de estimación convergentes a cero.

**CAPÍTULO VI: CONCLUSIONES, LIMITACIONES Y  
RECOMENDACIONES**

## CONCLUSIONES

- El pronóstico de las enfermedades más prevalentes mediante Redes Neuronales Artificiales en el Hospital Víctor Ramos Guardia de Huaraz, son adecuadas y con convergencia del error de estimación a cero, sin superar el umbral o límite de error de estimación del 1% con un tope máximo de 1000000 iteraciones.
- Las enfermedades más prevalentes en los pacientes del Hospital Víctor Ramos Guardia de Huaraz son: Factores que influyen en el estado de salud y contacto con los servicios de salud; Traumatismos, envenenamientos y algunas otras consecuencias de causa externa; Enfermedades del sistema respiratorio; Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte y Ciertas enfermedades infecciosas y parasitarias, que en conjunto representan al 79,1% del total de casos.
- La cantidad de neuronas de entrada para el diseño del modelo de red neuronal artificial para realizar el pronóstico de las enfermedades prevalentes en el Hospital Víctor Ramos Guardia de Huaraz, fue de cuatro neuronas, donde la primera neurona representa la serie de tiempo rezagada en un periodo, la segunda neurona representa la serie de tiempo rezagada en dos periodos, la tercera neurona representa la serie de tiempo rezagada en tres periodos y la cuarta neurona representa la serie de tiempo rezagada en cuatro periodos.
- El modelo de Redes Neuronales Artificiales adecuado para el pronóstico de las enfermedades prevalentes en el Hospital Víctor Ramos Guardia de Huaraz, presenta la siguiente arquitectura: red neuronal artificial modelo de perceptron

multicapa, con función de activación e integración logística, con método de retropropagación del error, con una capa de entrada con cuatro neuronas, dos capas ocultas (para la serie E1 compuesta por 8 y 6 neuronas, para la serie E2 compuesta por 6 y 6 neuronas, para la serie E3 compuesta por 7 y 7 neuronas, para la serie E4 compuesta por 4 y 4 neuronas y para la serie E5 compuesta por 5 y 6 neuronas) y una capa de salida con una neurona.

## **LIMITACIONES**

Las limitaciones encontradas son en cuanto al acceso total de las bases de datos del Hospital Víctor Ramos Guardia, ya que los datos de años pasados no se encuentran integradas a los nuevos sistemas de registro con que cuenta el hospital.

Limitación en cuanto a material bibliográfico, ya que en la biblioteca de la facultad de ciencias no existen libros especializados en cuanto a redes neuronales artificiales.

## **RECOMENDACIONES**

A los responsables del Área de Estadística del Hospital Víctor Ramos Guardia, para que implementen y enlacen las bases de datos a fin de utilizar toda la data de manera conjunta para un mejor análisis de las enfermedades en los pacientes del hospital, con la finalidad de que prioricen políticas de salud acordes a la realidad de la población.

Al director de escuela profesional de estadística e informática, que implementen cursos de redes neuronales artificiales, para que los estudiantes tengan las herramientas actuales y acordes al avance de la tecnología.

A los estudiantes de la Escuela Académico Profesional de Estadística e Informática, para que profundicen sus conocimientos en temas de redes neuronales artificiales para realizar pronósticos de series de tiempo.

## REFERENCIAS BIBLIOGRÁFICAS

**Asnate Salazar, Edwin Johny (Perú, 2014).** Factores asociados al bajo peso de los recién nacidos, del Hospital Víctor Ramos Guardia, Huaraz, 2002-2009, usando minería de datos con redes neuronales”.

**Arana, E. & Marti & Bonmatí, L. & Bautista, D. (España, 2003).** Diagnóstico de las lesiones de la calita. Selección de variables por redes neuronales y regresión logística, Neurocirugía, pág. 377-384.

**Bails, Dale & Larry Peppers (1993).** Business Fluctuations, Forecasting Techniques and Applications, Prentice-Hall.

URL:[http://148.206.107.15/biblioteca\\_digital/capitulos/382-5282doe.pdf](http://148.206.107.15/biblioteca_digital/capitulos/382-5282doe.pdf)

**Bernal López, Edwin Andrés (Colombia-Bogotá, 1999).** Redes Neuronales (Inteligencia Artificial) Aplicada al Diagnóstico Médico.

**Camilo, S. (Colombia, 2006).** Predicción de series temporales con redes neuronales: Una aplicación a la inflación, revista Colombiana de Estadística, pág. 77-92.

**Colina, E. & Rivas, F. (Venezuela, 1998).** Introducción a la inteligencia artificial. Mérida, Venezuela: Universidad de Los Andes. pág. 231-250.

**Collin, L (1989).** Epidemiología y Estadística

**Collantes Duarte, Joanna Verónica (Venezuela-Mérida, 2001).** Predicción con redes neuronales.

**Delgado Alberto (Colombia-Bogotá, 1999).** Aplicación de las Redes Neuronales Artificiales en Medicina.

**Faraway, J. & Chatfield, C. (1988).** Time series forecasting with neural networks: a comparative study using the airline data. Applied Statistic, pág. 231-250.

**Guzmán, M. & Carrillo, H. & Villaseñor, E. & Valencia, E. & Calero, R. & Moran, y otros (2004).** Minería de Datos con Redes Neuronales Artificiales: Aplicación en Vacunas - Tuberculosis. Centro de Investigación-Desarrollo y Producción de Vacunas y Sueros. Laboratorio de Dinámica no Lineal Facultad de Ciencias UNAM.

URL:<http://www.dynamics.unam.edu/DinamicaNoLineal2/Proyectos/ProyectoPAPIME.html>

**Hagan, M. & Demuth, H. & Beale, M (USA, 1996).** Neuronal network design. Boston, USA: PWS Publishing Company.

**Haykin, S. (USA, 1994).** Neural networks. a comprehensive foundation. USA: Macmillan College Publishing Company.

**Hernández Orallo, J. & Ramírez Quintana, C. & Ramírez, F. (2004).** Introducción a la minería de datos. Editorial Pearson.

**Hernández, Miguel (2013).** Redes Neuronales, técnicas estadísticas en análisis de mercados.

**Hilera, José & Martínez, Víctor (España, 1955).** Redes neuronales artificiales, fundamentos modelos y aplicaciones, Rama Editorial.

URL:[https://pirhua.udep.edu.pe/bitstream/handle/123456789/1745/ING\\_533.pdf?sequence=1](https://pirhua.udep.edu.pe/bitstream/handle/123456789/1745/ING_533.pdf?sequence=1)

**Hill, T. O'Connor & Remus, W. (1996). Neural network models for time series forecast management science. 42(7), pág. 1082-1092.**

URL:<http://pisis.fime.uanl.mx/ftp/pubs/reports/2007/pisis-2007-01.pdf>

**MPérez, B.A. & Fernández,R. & Wright, E.R. & Ascitutto, K. & Rosato, M.;Lahore, S. & Romero, A.M. & Babbitt, S. E Irigoyen, E. (2005), pág. 230.**  
Fitopatólogos.

**Makridakis, S. & Wheelwright, S. (1992).** Manual de técnicas de pronósticos.  
México: Limusa Noriega Editores.

**Nojek, S. & Britos, P. & Rossi, B. & García Martínez, R. (Argentina, 2003).**  
Pronóstico de Ventas: Comparación de Predicción basada en Redes Neuronales versus Método Estadístico. Reportes Técnicos en Ingeniería del Software.

**Pérez, Ma. Luisa & Quintín, Martín (España, 2003).** Aplicaciones de las redes neuronales artificiales a la estadística: La Muralla.

URL:<http://tesis.ipn.mx/jspui/bitstream/123456789/9615/1/251.pdf>

**Pérez & Martín (2003).** Unidades de Procesamiento de una Red Neuronal Artificial.

**Salud Pública y Epidemiológica (2014).** Epidemiología

**Trujillano Cabello, J (España, 2005).** Redes Neuronales Artificiales en Medicina Intensiva.

**Valenzuela, M (2008).** Logística Inteligente. Centro de Computación Inteligente y Robótica Tecnológico de Monterrey, Campus Monterrey. Disponible en URL:[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)

**Vásquez, S (2010).** Tecnología de Informática(2010)

URL:<https://solvasquez.wordpress.com/2010/08/15/inteligencia-artificial/>

**Viñuela, Pedro Isasi & Galván, León, Inés M. (Madrid, 2003).** Redes de Neuronas Artificiales, un enfoque práctico.

**Wedding & Cios (1996).** Función de base radial.

**Wong, F. (1991).** Time series forecasting using backpropagation neural networks. Neurocomputing, pág. 147-159

URL:<http://www.wseas.us/elibrary/conferences/switzerland2002/papers/464.pdf>

## **ANEXOS**

Nº FORMATO  
CÓD. DEL ESTADÍSTICO

# MINISTERIO DE SALUD

## OFICINA GENERAL DE ESTADÍSTICA E INFORMÁTICA

### Registro Diario de Atención y Otras Actividades

1	
TURNO	
MAÑANA	TARDE

2		3				4		5					
FECHA		DPTO	PROV	DIST	ESTAB	SERVICIO		NOMBRE DEL RESPONSABLE DE LA ATENCIÓN					
MES	AÑO												
6	7	8	9	10	11	12	13			14		15	16
DIA	H.C./FICHA FAMILIAR	DISTRITO PROCED.	EDAD	SEXO	ESTAB	SERV.	DIAGNÓSTICO, MOTIVO DE LA CONSULTA Y/O ACTIVIDAD DE SALUD			TIPO DE DIAGNÓSTICO		LAB	CÓDIG
				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
2				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
3				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
4				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
5				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
6				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
7				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
8				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
9				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
10				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
11				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	
12				M	N	N				P	D	R	
				F	C	C				P	D	R	
					R	R				P	D	R	

ÍTEM 10 SEXO: M - MASCULINO; F - FEMENINO      ÍTEM 11 (ESTAB) Y 12 (SERV): N - NUEVO; C - CONTINUADOR; R - REINGRESANTE  
 ÍTEM 14 (TIPO DE DIAGNÓSTICO): P - PRESUNTIVO; D - DEFINITIVO; R - REPETIDO  
 LA HOJA DEBE SER LLENADA INTEGRAMENTE POR EL PRESTADOR DE SALUD. LOS RECUADROS SOMBRADOS SERÁN DE USO DEL ESTADÍSTICO.

## Anexo 2

### *Prevalencia de las enfermedades de los pacientes del Hospital Víctor Ramos*

*Guardia, 2011 – 2016.*

<b>Enfermedad</b>	<b>Número de pacientes</b>	<b>Porcentaje</b>
Factores que influyen en el estado de salud y contacto con los servicios de salud	45331	24,3%
Traumatismos, envenenamientos y algunas otras consecuencias de causa externa	39054	21,0%
Enfermedades del sistema respiratorio	24005	12,9%
Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte	23074	12,4%
Ciertas enfermedades infecciosas y parasitarias	15857	8,5%
Enfermedades del aparato digestivo	8750	4,7%
Enfermedades del aparato genitourinario	5919	3,2%
Trastornos mentales y del comportamiento	4474	2,4%
Enfermedades del sistema circulatorio	3887	2,1%
Enfermedades de la piel y el tejido subcutáneo	3585	1,9%
Enfermedades del sistema osteomuscular y del tejido conectivo	3083	1,7%
Embarazo, parto y puerperio	2517	1,4%
Enfermedades del sistema nervioso	1769	0,9%
Enfermedades del oído y de la apófisis mastoides	1378	0,7%
Enfermedades endocrinas, nutricionales y metabólicas	1134	0,6%
Neoplasias	594	0,3%
Ciertas afecciones originadas en el periodo perinatal	558	0,3%
Causas externas de morbilidad y de mortalidad	524	0,3%
Enfermedades del ojo y sus anexos	379	0,2%
Enfermedades de la sangre y de los órganos hematopoyéticos y otros trastornos que afectan el mecanismo de la inmunidad	307	0,2%
Malformaciones congénitas, deformidades y anomalías cromosómicas	114	0,1%
<b>Total</b>	<b>186293</b>	<b>100,0%</b>

Fuente: Elaboración propia

Anexo 3

*Enfermedades según sexo de los pacientes del Hospital Víctor Ramos Guardia,  
2011 – 2016*

Sexo	Factores que influyen en el estado de salud y contacto con los servicios de salud		Traumatismos, envenenamientos y algunas otras consecuencias de causa externa		Enfermedades del sistema respiratorio		Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte		Ciertas enfermedades infecciosas y parasitarias		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Masculino	9012	19,9%	25475	65,2%	12441	51,8%	10287	44,6%	7588	47,9%	<b>64803</b>	<b>44,0%</b>
Femenino	36319	80,1%	13579	34,8%	11564	48,2%	12787	55,4%	8269	52,1%	<b>82518</b>	<b>56,0%</b>
<b>Total</b>	<b>45331</b>	<b>100,0%</b>	<b>39054</b>	<b>100,0%</b>	<b>24005</b>	<b>100,0%</b>	<b>23074</b>	<b>100,0%</b>	<b>15857</b>	<b>100,0%</b>	<b>147321</b>	<b>100,0%</b>

Fuente: Elaboración propia

Anexo 4

*Enfermedades según edad de los pacientes del Hospital Víctor Ramos Guardia, 2011 – 2016*

Edad	Factores que influyen en el estado de salud y contacto con los servicios de salud		Traumatismos, envenenamientos y algunas otras consecuencias de causa externa		Enfermedades del sistema respiratorio		Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte		Ciertas enfermedades infecciosas y parasitarias		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
<= 4	3266	7,2%	4539	11,6%	11969	49,9%	5613	24,3%	5113	32,2%	<b>30500</b>	<b>20,7%</b>
5 - 9	1146	2,5%	3202	8,2%	2716	11,3%	1648	7,1%	946	6,0%	<b>9658</b>	<b>6,6%</b>
10 - 14	1011	2,2%	2587	6,6%	1226	5,1%	1278	5,5%	669	4,2%	<b>6771</b>	<b>4,6%</b>
15 - 19	5502	12,1%	3615	9,3%	944	3,9%	1580	6,8%	929	5,9%	<b>12570</b>	<b>8,5%</b>
20 - 24	9964	22,0%	4751	12,2%	1164	4,8%	1855	8,0%	1313	8,3%	<b>19047</b>	<b>12,9%</b>
25 - 29	8275	18,3%	4168	10,7%	1007	4,2%	1707	7,4%	1169	7,4%	<b>16326</b>	<b>11,1%</b>
30 - 34	5780	12,8%	3252	8,3%	768	3,2%	1393	6,0%	928	5,9%	<b>12121</b>	<b>8,2%</b>
35 - 39	3875	8,5%	2634	6,7%	648	2,7%	1290	5,6%	754	4,8%	<b>9201</b>	<b>6,2%</b>
40 - 44	1798	4,0%	2120	5,4%	536	2,2%	1034	4,5%	602	3,8%	<b>6090</b>	<b>4,1%</b>
45 - 49	852	1,9%	1679	4,3%	468	1,9%	903	3,9%	594	3,7%	<b>4496</b>	<b>3,1%</b>
50 - 54	758	1,7%	1491	3,8%	490	2,0%	873	3,8%	639	4,0%	<b>4251</b>	<b>2,9%</b>
55 - 59	651	1,4%	1181	3,0%	373	1,6%	804	3,5%	464	2,9%	<b>3473</b>	<b>2,4%</b>
60 - 64	529	1,2%	1012	2,6%	306	1,3%	672	2,9%	439	2,8%	<b>2958</b>	<b>2,0%</b>
65 - 69	488	1,1%	848	2,2%	298	1,2%	636	2,8%	366	2,3%	<b>2636</b>	<b>1,8%</b>
70 - 74	419	0,9%	691	1,8%	270	1,1%	527	2,3%	325	2,0%	<b>2232</b>	<b>1,5%</b>
75 - 79	424	0,9%	513	1,3%	275	1,1%	552	2,4%	251	1,6%	<b>2015</b>	<b>1,4%</b>
80+	593	1,3%	771	2,0%	547	2,3%	709	3,1%	356	2,2%	<b>2976</b>	<b>2,0%</b>
<b>Total</b>	<b>45331</b>	<b>100,0%</b>	<b>39054</b>	<b>100,0%</b>	<b>24005</b>	<b>100,0%</b>	<b>23074</b>	<b>100,0%</b>	<b>15857</b>	<b>100,0%</b>	<b>147321</b>	<b>100,0%</b>

Fuente: Elaboración propia

Anexo 5

*Enfermedades según tipo de paciente del Hospital Víctor Ramos Guardia, 2011 – 2016*

Tipo de Paciente	Factores que influyen en el estado de salud y contacto con los servicios de salud		Traumatismos, envenenamientos y algunas otras consecuencias de causa externa		Enfermedades del sistema respiratorio		Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte		Ciertas enfermedades infecciosas y parasitarias		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Sin seguro	44863	99,0%	38773	99,3%	23799	99,1%	22841	99,0%	15744	99,3%	<b>146020</b>	<b>99,1%</b>
SIS	322	0,7%	133	0,3%	105	0,4%	144	0,6%	56	0,4%	<b>760</b>	<b>0,5%</b>
EsSalud	24	0,1%	14	0,0%	11	0,0%	9	0,0%	5	0,0%	<b>63</b>	<b>0,0%</b>
Seguro militar	14	0,0%	17	0,0%	14	0,1%	8	0,0%	9	0,1%	<b>62</b>	<b>0,0%</b>
Seguro privado	62	0,1%	44	0,1%	35	0,1%	27	0,1%	13	0,1%	<b>181</b>	<b>0,1%</b>
SOAT	23	0,1%	38	0,1%	18	0,1%	14	0,1%	6	0,0%	<b>99</b>	<b>0,1%</b>
Otro	23	0,1%	35	0,1%	23	0,1%	31	0,1%	24	0,2%	<b>136</b>	<b>0,1%</b>
<b>Total</b>	<b>45331</b>	<b>100,0%</b>	<b>39054</b>	<b>100,0%</b>	<b>24005</b>	<b>100,0%</b>	<b>23074</b>	<b>100,0%</b>	<b>15857</b>	<b>100,0%</b>	<b>147321</b>	<b>100,0%</b>

Fuente: Elaboración propia

Tabla 6

*Enfermedades según motivo de atención de los pacientes del Hospital Víctor Ramos*

*Guardia, 2011 – 2016*

Motivo de Atención	Factores que influyen en el estado de salud y contacto con los servicios de salud		Traumatismos, envenenamientos y algunas otras consecuencias de causa externa		Enfermedades del sistema respiratorio		Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte		Ciertas enfermedades infecciosas y parasitarias		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Agresión	251	0,6%	769	2,0%	5	0,0%	6	0,0%	5	0,0%	<b>1036</b>	<b>0,7%</b>
Accidente de tránsito	787	1,7%	2782	7,1%	4	0,0%	17	0,1%	3	0,0%	<b>3593</b>	<b>2,4%</b>
Accidente de trabajo	34	0,1%	72	0,2%	3	0,0%	1	0,0%	0	0,0%	<b>110</b>	<b>0,1%</b>
Intento de suicidio	17	0,0%	107	0,3%	3	0,0%	4	0,0%	3	0,0%	<b>134</b>	<b>0,1%</b>
Enfermedad súbita	4599	10,1%	30934	79,2%	86	0,4%	553	2,4%	61	0,4%	<b>36233</b>	<b>24,6%</b>
Enfermedad común	39254	86,6%	3864	9,9%	23463	97,7%	22265	96,5%	15546	98,0%	<b>104392</b>	<b>70,9%</b>
Violación	1	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	<b>1</b>	<b>0,0%</b>
Ignorado	358	0,8%	344	0,9%	434	1,8%	216	0,9%	122	0,8%	<b>1474</b>	<b>1,0%</b>
Intoxicación	30	0,1%	182	0,5%	7	0,0%	12	0,1%	117	0,7%	<b>348</b>	<b>0,2%</b>
<b>Total</b>	<b>45331</b>	<b>100,0%</b>	<b>39054</b>	<b>100,0%</b>	<b>24005</b>	<b>100,0%</b>	<b>23074</b>	<b>100,0%</b>	<b>15857</b>	<b>100,0%</b>	<b>147321</b>	<b>100,0%</b>

Fuente: Elaboración propia

Anexo 7

*Enfermedades según servicio de atención de los pacientes del Hospital Víctor*

*Ramos Guardia, 2011 – 2016*

Servicio donde se atiende	Factores que influyen en el estado de salud y contacto con los servicios de salud		Traumatismos, envenenamientos y algunas otras consecuencias de causa externa		Enfermedades del sistema respiratorio		Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte		Ciertas enfermedades infecciosas y parasitarias		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Medicina general	7961	17,6%	2218	5,7%	8381	34,9%	13838	60,0%	9222	58,2%	<b>41620</b>	<b>28,3%</b>
Neumología	0	0,0%	2	0,0%	1	0,0%	6	0,0%	1	0,0%	<b>10</b>	<b>0,0%</b>
Cardiología	3	0,0%	21	0,1%	3	0,0%	8	0,0%	4	0,0%	<b>39</b>	<b>0,0%</b>
Neurología	0	0,0%	2	0,0%	0	0,0%	0	0,0%	1	0,0%	<b>3</b>	<b>0,0%</b>
Gastroenterología	7	0,0%	1	0,0%	0	0,0%	1	0,0%	0	0,0%	<b>9</b>	<b>0,0%</b>
Dermatología	0	0,0%	1	0,0%	0	0,0%	0	0,0%	0	0,0%	<b>1</b>	<b>0,0%</b>
Psiquiatría	23	0,1%	4	0,0%	58	0,2%	36	0,2%	24	0,2%	<b>145</b>	<b>0,1%</b>
Otros de medicina	7	0,0%	1	0,0%	1	0,0%	4	0,0%	0	0,0%	<b>13</b>	<b>0,0%</b>
Cirugía	6028	13,3%	33545	85,9%	53	0,2%	1135	4,9%	53	0,3%	<b>40814</b>	<b>27,7%</b>
Traumatología/ortopedia	46	0,1%	1531	3,9%	1	0,0%	6	0,0%	1	0,0%	<b>1585</b>	<b>1,1%</b>
Otorrinolaringología	8	0,0%	58	0,1%	14	0,1%	33	0,1%	0	0,0%	<b>113</b>	<b>0,1%</b>
Oftalmología	8	0,0%	134	0,3%	0	0,0%	0	0,0%	0	0,0%	<b>142</b>	<b>0,1%</b>
Urología	0	0,0%	0	0,0%	0	0,0%	2	0,0%	0	0,0%	<b>2</b>	<b>0,0%</b>
Neurocirugía	7	0,0%	140	0,4%	0	0,0%	23	0,1%	4	0,0%	<b>174</b>	<b>0,1%</b>
Otros de cirugía	0	0,0%	26	0,1%	0	0,0%	0	0,0%	1	0,0%	<b>27</b>	<b>0,0%</b>
Pediatría general	3642	8,0%	1115	2,9%	15467	64,4%	7761	33,6%	6507	41,0%	<b>34492</b>	<b>23,4%</b>
Obstetricia	26801	59,1%	141	0,4%	22	0,1%	84	0,4%	23	0,1%	<b>27071</b>	<b>18,4%</b>
Ginecología	782	1,7%	88	0,2%	4	0,0%	135	0,6%	16	0,1%	<b>1025</b>	<b>0,7%</b>
Oncología	2	0,0%	0	0,0%	0	0,0%	1	0,0%	0	0,0%	<b>3</b>	<b>0,0%</b>
Clínica cirugía	6	0,0%	25	0,1%	0	0,0%	1	0,0%	0	0,0%	<b>32</b>	<b>0,0%</b>
Unidad de cuidados intensivos	0	0,0%	1	0,0%	0	0,0%	0	0,0%	0	0,0%	<b>1</b>	<b>0,0%</b>
<b>Total</b>	<b>45331</b>	<b>100,0%</b>	<b>39054</b>	<b>100,0%</b>	<b>24005</b>	<b>100,0%</b>	<b>23074</b>	<b>100,0%</b>	<b>15857</b>	<b>100,0%</b>	<b>147321</b>	<b>100,0%</b>

Fuente: Elaboración propia

Anexo 8

*Enfermedades según destino del paciente post servicio del Hospital Víctor Ramos*

*Guardia, 2011 – 2016*

Destino del paciente	Factores que influyen en el estado de salud y contacto con los servicios de salud		Traumatismos, envenenamientos y algunas otras consecuencias de causa externa		Enfermedades del sistema respiratorio		Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte		Ciertas enfermedades infecciosas y parasitarias		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Domicilio	18368	40,5%	31242	80,0%	21753	90,6%	19089	82,7%	14203	89,6%	<b>104655</b>	<b>71,0%</b>
Hospitalización	16549	36,5%	1951	5,0%	1530	6,4%	1451	6,3%	912	5,8%	<b>22393</b>	<b>15,2%</b>
Transferencia	18	0,0%	15	0,0%	11	0,0%	76	0,3%	5	0,0%	<b>125</b>	<b>0,1%</b>
Referencia	517	1,1%	249	0,6%	52	0,2%	74	0,3%	21	0,1%	<b>913</b>	<b>0,6%</b>
Fuga	611	1,3%	470	1,2%	95	0,4%	277	1,2%	99	0,6%	<b>1552</b>	<b>1,1%</b>
Defunción	20	0,0%	16	0,0%	14	0,1%	171	0,7%	6	0,0%	<b>227</b>	<b>0,2%</b>
Reposo	151	0,3%	3481	8,9%	153	0,6%	1041	4,5%	335	2,1%	<b>5161</b>	<b>3,5%</b>
Retiro Voluntario	388	0,9%	710	1,8%	93	0,4%	290	1,3%	63	0,4%	<b>1544</b>	<b>1,0%</b>
Ignorado	8709	19,2%	920	2,4%	304	1,3%	605	2,6%	213	1,3%	<b>10751</b>	<b>7,3%</b>
<b>Total</b>	<b>45331</b>	<b>100,0%</b>	<b>39054</b>	<b>100,0%</b>	<b>24005</b>	<b>100,0%</b>	<b>23074</b>	<b>100,0%</b>	<b>15857</b>	<b>100,0%</b>	<b>147321</b>	<b>100,0%</b>

Fuente: Elaboración propia