

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INFORMÁTICA

**USO DE TÉCNICAS DE INTELIGENCIA COMPUTACIONAL
PARA LA ESTIMACIÓN DE LESIONADOS Y FALLECIDOS EN
ACCIDENTES DE TRANSITO**

JAIME ENRIQUE PÉREZ VARAS

TESIS DE GRADO
MAGÍSTER EN INGENIERÍA INFORMÁTICA

DICIEMBRE, 2013

Pontificia Universidad Católica de Valparaíso
Facultad de Ingeniería
Escuela de Ingeniería Informática

**USO DE TÉCNICAS DE INTELIGENCIA COMPUTACIONAL
PARA LA ESTIMACIÓN DE LESIONADOS Y FALLECIDOS EN
ACCIDENTES DE TRANSITO**

JAIME ENRIQUE PÉREZ VARAS

Director de Tesis: **José Miguel Rubio León**

Programa: **Magíster en Ingeniería Informática**

DICIEMBRE, 2013

Agradecimientos

A mis padres y hermana, por todo el apoyo brindado durante toda mi vida, por todos sus esfuerzos, por creer en mis capacidades y alentarme cada día.

A mi profesor guía por pensar en incorporarme a este proyecto, por sus revisiones, por otorgarme el apoyo y la comprensión necesaria para poder terminar este proyecto.

A mi profesor Co-referente por aportar con su conocimiento, apoyo y correcciones.

A todos los compañeros del Magíster, en especial a mis grupos de trabajo ya que sin ellos esta etapa no sería posible.

En especial a la Familia Muñoz Álvarez, por darme todo el apoyo en este camino ya que sin su apoyo esto no sería posible.

Especialmente a mi novia, compañera y amiga Diana, por su grandiosa ayuda y apoyo. Por dejarme sacrificar nuestros fines de semanas y acompañarme en cada paso que doy en esta vida.

Resumen

Dentro de la Inteligencia Computacional unos de sus principales paradigmas son las redes neuronales artificiales.

En este contexto, las redes neuronales artificiales son un paradigma de aprendizaje y procesamiento automático, con años de estudio (desde 1943), pero su utilización en problemas de estimación de accidentes es reciente ya que se han analizado pero con técnicas de minería de datos, la utilización de computación evolutiva no han sido incorporadas del todo dentro de estas investigaciones, pese a que son alternativas válidas para obtener resultados satisfactorios.

Debido a lo anterior, en este proyecto se propone implementar modelos de estimación del número de personas lesionadas y fallecidas en accidentes de tránsito en la región de Valparaíso, Chile, utilizando redes neuronales artificiales en conjunto con algoritmos de optimización por enjambre de partículas.

Palabras Clave: Inteligencia Computacional, Redes Neuronales Artificiales , Optimización por enjambre de partículas (PSO), Accidentes de Tránsito.

Abstract

Computational Intelligence In one of its main paradigms are artificial neural networks.

In this context, artificial neural networks are a paradigm of learning and automatic processing, with years of study (since 1943), but their use in estimation problems is recent accidents have been analyzed but data mining techniques, the use of evolutionary computation has not been fully incorporated into these investigations, although they are valid alternatives to obtain satisfactory results.

Due to the above, this project aims to implement models to estimate the number of dead and injured in traffic accidents in the region of Valparaíso, Chile, using artificial neural network algorithms in conjunction with particle swarm optimization.

Keywords: Computational Intelligence, Artificial Neural Networks, Particle Swarm Optimization (PSO), traffic accidents.

Índice

Resumen	I
Indice	IV
Lista de Abreviaturas	V
Lista de Figuras	VI
Lista de Tablas	VII
1. Introducción	1
2. Definición de Objetivos	3
2.1. Objetivo General	3
2.2. Objetivos Específicos	3
3. Metodología	4
3.1. Entrenamiento de la red neuronal	4
3.2. Métrica de validación	5
3.3. Uso de bases de datos	6
3.4. Diseño de Modelos	7
4. Plan de Trabajo	8
5. Marco Teórico	9
5.1. Inteligencia Computacional	9

5.2.	Redes Neuronales Artificiales	9
5.2.1.	Reseña Histórica	11
5.2.2.	Fundamentos Biológicos	14
5.2.3.	La Neurona Biológica	14
5.2.4.	La Neurona Artificial	18
5.2.5.	Clasificación de Redes Neuronales	20
5.2.6.	Estructura de una Red	21
5.2.7.	Aprendizaje	23
5.2.7.1.	Aprendizaje Supervisado	25
5.2.7.2.	Aprendizaje No Supervisado	26
5.2.7.3.	Aprendizaje Híbrido	27
5.3.	Optimización por Enjambre de Partículas	27
5.3.1.	Algoritmo PSO Modelo Tradicional	28
5.3.2.	Ajuste de los Parámetros de PSO	32
5.3.3.	PSO con Búsqueda Local Adaptativa (LSPSO)	36
5.4.	Análisis de los Accidentes de Tránsito	37
6.	Desarrollo del Modelo	41
6.1.	Graficos de representación de datos	43
6.2.	Desarrollo del Modelo	44
6.3.	Métricas de Calidad	46
6.3.1.	Raíz Error Cuadrático Medio (RMSE)	46
6.3.2.	Coefficiente de Determinación (R^2)	46
6.4.	Entorno de Trabajo	46

7. Resultados Computacionales	47
7.1. Análisis de los resultados con Modelo Tradicional PSO - Fallecidos del Gran Valparaíso	47
7.2. Análisis de los resultados con Modelo Tradicional PSO - Lesionados del Gran Valparaíso	51
7.3. Análisis de los resultados con Modelo Tradicional PSO - Muertos Otras Comunas	54
7.4. Análisis de los resultados con Modelo Tradicional PSO - Lesionados Otras Comunas	57
7.5. Análisis de los resultados con Modelo PSO con Búsqueda Local Adaptativa - Fallecidos del Gran Valparaíso	61
7.6. Análisis de los resultados con Modelo PSO con Búsqueda Local Adaptativa - Lesionados del Gran Valparaíso	64
7.7. Análisis de los resultados con Modelo PSO con Búsqueda Local Adaptativa - Fallecidos Otras Comunas	68
7.8. Análisis de los resultados con Modelo PSO con Búsqueda Local Adaptativa - Lesionados Otras Comunas	71
7.9. Consolidación de resultados obtenidos	74
8. Conclusión	76

Lista de Abreviaturas

OMS	Organización Mundial de la Salud
CONASET	Comisión Nacional de Seguridad de Tránsito
OPS	Organización Panamericana de la Salud
RNA	Redes Neuronales Artificiales
PSO	Optimización por enjambre de partículas
IC	Inteligencia Computacional
GOFAI	Good Old-Fashioned Artificial Intelligence
ADALINE	Adaptative Linear Element
RBF	Radial Basis Function
BP	Back Propagation
MLP	feed-forward

Lista de Figuras

1.	Neurona biológica	15
2.	Neurona artificial	19
3.	Red neuronal artificial de tres capas	22
4.	Aprendizaje supervisado	25
5.	Aprendizaje no supervisado	27
6.	Enjambres en la Naturaleza	28
7.	Entidades base de datos CONASET	42
8.	Cantidad de Fallecidos	43
9.	Cantidad de lesionados	44
10.	Grafico 16 nodos de entrada, 7 ocultos	50
11.	Grafico 12 nodos de entrada, 3 ocultos	53
12.	Grafico 13 nodos de entrada, 3 ocultos	57
13.	Grafico 18 nodos de entrada, 2 ocultos	60
14.	Grafico 11 nodos de entrada, 3 ocultos	64
15.	Grafico 13 nodos de entrada, 3 ocultos	67
16.	Grafico 11 nodos de entrada, 12 ocultos	70
17.	Grafico 11 nodos de entrada, 2 ocultos	74

Lista de Tablas

1.	Plan de Trabajo.	8
2.	PSO - Muertos Gran Valparaíso	49
3.	PSO - Lesionados Gran Valparaíso	52
4.	PSO - Muertos Otras Comunas	56
5.	PSO - Lesionados Otras Comunas	59
6.	LPSO - Muertos Gran Valparaíso	63
7.	LPSO - Lesionados Gran Valparaíso	66
8.	LPSO - Fallecidos Otras Comunas	69
9.	LPSO - Lesionados Otras Comunas	73
10.	Resumen de Resultados.	75

1. Introducción

Como se indica en el sitio web de la organización mundial de la salud (OMS), por medio de su oficina regional (OPS) en su informe sobre la situación mundial de la seguridad vial señala que: “En las Américas y el Caribe, cada año fallecieron alrededor de 130.000 personas, más de 1,2 millones sufrieron traumatismos, y cientos de miles resultan discapacitadas como consecuencia de colisiones, choques, volcaduras o atropellamientos en las vías de tránsito”.

En Chile la situación no es diferente. Los costos por siniestros de tránsito solamente en el año 2010 representa el 1,5% del PIB nacional, esto considerando los costos sociales y privados. Los accidentes de tránsito se han posicionado como una verdadera epidemia, que ha llevado a alcanzar altas cifras de personas lesionadas y fallecidas. Esta situación se observa como un problema complejo de solucionar, según la Comisión Nacional de Seguridad de Tránsito (CONASET), en 2011 se contabilizaron un total de 62.834 accidentes, siendo la colisión el tipo de siniestro más recurrente. La región metropolitana se transforma en la zona con más accidentes a nivel nacional y el 2011 contribuyó con un total de 23.520 siniestros, con un total de 360 personas fallecidas y 15.549 en la categoría de lesionadas; sufriendo lesiones graves 2.008 personas, menos graves 1.200 personas y leves 12.341 personas. Las regiones de Valparaíso y del Biobío siguen a la región metropolitana en este trágico ranking, con 7.724 y 6.643 siniestros respectivamente [5].

Se han propuesto modelos predictivos que permiten anticiparse a las causas de éste tipo de incidentes, en los cuales se han obtenido resultados que hacen avanzar en éste ámbito, como por ejemplo [3]. Técnicas como la Minería de Datos (Data Mining), permite visualizar patrones y variables significativas que ayudan a identificar de mejor manera las condiciones en las que ocurren los accidentes, tomando en cuenta que éste tipo de evento contempla un sin número de variables, de las cuales se deben considerar sólo las más relevantes y así poder cuantificar la cantidad de accidentes que se producen en la actualidad.

La Minería de Datos cuenta con diversas técnicas que permiten el análisis de datos, dentro de las cuales se puede mencionar las Redes Neuronales Artificiales (RNA), que corresponden a un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los animales. Estos mecanismos son utilizados para la clasificación y predicción de datos, y en combinación con algoritmos de computación evolutiva (optimización por enjambre de partículas (PSO)), se pueden establecer patrones que permiten estimar el número de personas lesionadas y fallecidas en accidentes de tránsito.

Se ha estructurado el presente informe de avance otorgándoles al lector en primera instancia los objetivos del proyecto, la metodología, seguido de su plan de trabajo. Seguido se hace un estado del arte de los tópicos mas importantes a tratar en esta investigación, terminando con una conclusión del presente informe

2. Definición de Objetivos

2.1. Objetivo General

Desarrollar modelos para estimar el número de personas lesionadas y fallecidas en accidentes de tránsito de la región de Valparaíso (Chile), utilizando redes neuronales artificiales en combinación con algoritmos de computación evolutiva.

2.2. Objetivos Específicos

- Diseñar la estructura del modelo de estimación neuronal de personas lesionadas y fallecidas en accidentes de tránsito de la región de Valparaíso (Chile) y ajustar sus parámetros usando PSO.
- Implementar modelos de estimación de personas lesionadas y fallecidas en accidentes de tránsito de la región de Valparaíso (Chile), basados en redes neuronales con PSO y LPSO.
- Evaluar y contrastar el rendimiento de los modelos de estimación neuronal implementados.

3. Metodología

Para resolver el problema propuesto se ha planteado la siguiente metodología, con el análisis de los diversos tópicos que se explican a continuación.

3.1. Entrenamiento de la red neuronal

El aprendizaje en las redes neuronales artificiales es visto como el proceso de ajuste de parámetros libres. Se parte de un conjunto de pesos sinápticos aleatorios, con el fin de buscar un conjunto de pesos que permitan a la red desarrollar correctamente una tarea en particular. El proceso de aprendizaje es un proceso iterativo, en el cual se va refinando la solución hasta alcanzar un nivel de operación que de conformidad.

Una gran cantidad de métodos de entrenamiento utilizados en las redes neuronales con conexión hacia delante consisten en proponer una función de error que mida el rendimiento de la red en función de los pesos sinápticos. El objetivo que cumple el entrenar la red es obtener un conjunto de pesos sinápticos que minimizan (o maximizan) la función. El método de optimización proporciona una regla de actualización de los pesos que en función de los patrones de entrada modifica iterativamente los pesos hasta alcanzar el punto óptimo de la red neuronal. Una forma de obtener mejores resultado en la obtención de estos pesos, es mediante la utilización de PSO, el cual será utilizados en este proyecto.

El backpropagation (retro-propagación) es el método de entrenamiento más

utilizado en redes con conexión hacia delante. Este es un método de aprendizaje supervisado, en el que se aplica un patrón de entrada, el cual se propaga por las distintas capas que componen la red hasta producir la salida de la misma. La salida se compara con la salida deseada y se calcula el error cometido por cada neurona de salida. Estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de las capas intermedias. Cada neurona recibe un error que es proporcional a su contribución sobre el error total de la red. Basándose en el error recibido, se ajustan los errores de los pesos sinápticos de cada neurona. Por lo general una forma de entrenar redes neuronales es mediante el error cuadrático medio, este es el método utilizado mayormente para fines predictivos y esta dado por la ecuación 1:

$$E = \frac{1}{N} \times \sum_{i=1}^N (d_i(n) - y_i(n))^2 \quad (1)$$

Donde N es el número de patrones para entrenar la red, d_i es el valor deseado para el i -ésimo patrón de entrenamiento e y_i es el valor que entrega la red para el i -ésimo patrón de entrenamiento.

3.2. Métrica de validación

La validación de los modelos propuestos se realizará mediante la técnica de Validación Cruzada. Este método consiste en dividir los datos en dos partes, una para pruebas y otra para entrenamiento de la red neuronal, un ejemplo es

usar el 80 % de los datos para entrenamiento y un 20 % para validación. Luego, los datos de entrenamiento se dividen en k partes, donde se va dejando un conjunto de datos para hacer entrenamiento y otro para pruebas. Finalmente, se calcula el promedio de los fitness de cada parte, este fitness será utilizado como medida del entrenamiento. Además se guardarán los valores de los pesos que den mejores resultados en el entrenamiento.

3.3. Uso de bases de datos

Para el estudio se utilizarán los datos suministrados por CONASET, estos datos se encuentran almacenados en planillas Excel y están disponibles de forma gratuita en la Web. Los datos de cada accidente se obtienen mediante un registro en plantillas (SIEC2), siguiendo un instructivo, para posteriormente ser ingresados en una base de datos de CONASET.

Los datos se pueden separar en tres tipos de entidades, cada una con sus respectivos atributos para fines de registro:

1. Accidentes: identificador del accidente, fecha, hora de accidente, comuna, causas, tipo de accidente, estado atmosférico, condición calzada, tipo de calzada, estado calzada, urbano/rural, y la información que indica la cantidad de ilesos, leves, graves, menos graves y muertos.
2. Persona: identificador del accidente, sexo, calidad (peatón, conductor, pasajero), edad, y resultado (muerto, grave, menos grave, leve, ileso).
3. Vehículo: identificador del accidente, servicio, y el tipo de vehículo.

En cada accidente puede haber una o muchas personas involucradas y uno o muchos vehículos involucrados.

3.4. Diseño de Modelos

Para llegar a la obtención de los modelos se ha de seguir una serie de pasos, estos se detallan a continuación:

1. Tratamiento de datos: en este paso se realiza un pre-procesamiento de los datos, en el cual se eliminan redundancias e inconsistencias para luego normalizar los datos para facilitar las tareas posteriores.
2. Elección de la red neuronal artificial y algoritmo de entrenamiento: en esta etapa se seleccionan el tipo de red usada para cada modelo.
 - a) *Selección de algoritmo de entrenamiento* (PSO o alguna de sus variantes).
 - b) *Elección de topologías*: se define el número de nodos de entrada de la red, el número de nodos de la capa oculta y de la capa de salida.
3. Entrenamiento de la red: la red se entrena mediante PSO, de manera de obtener los pesos de la red. Como se indicó anteriormente, el entrenamiento se realiza utilizando un 80 % de los datos.
4. Testeo de la red: se obtienen los resultados de la estimación (número de accidentes y número de personas fallecidas y lesionadas), para ello se utilizará el 20 % de los datos restantes que no se utilizaron para entrenamiento.
5. Comparación de resultados: en la etapa final se compararán los resultados obtenidos por los diversos modelos.

4. Plan de Trabajo

Como se puede apreciar en la tabla 1, el presente plan se estructura de la siguiente forma:

Actividad	Etapa	Objetivo Asociado	Fecha Ejecución
<ul style="list-style-type: none"> ★ Seleccionar las variables más relevantes relacionadas con los accidentes de tránsito. ★ Calibrar el número de nodos óptimos de la capa oculta del estimador neuronal (EN). ★ Representación de los parámetros (pesos) del EN mediante PSO. 	Diseño del modelo de estimación neuronal	Diseñar la estructura del modelo de estimación neuronal de personas lesionadas y fallecidas en accidentes de tránsito de la región de Valparaíso (Chile) y ajustar sus parámetros usando PSO.	Marzo a Junio 2013
<ul style="list-style-type: none"> ★ Implementar el EN con PSO (EN+PSO). ★ Evaluación del rendimiento de EN+PSO. 	Calibración del modelo EN+PSO	<ul style="list-style-type: none"> ■ Implementar modelos de estimación de personas lesionadas y fallecidas en accidentes de tránsito de la región de Valparaíso (Chile), basados en redes neuronales con PSO. ■ Evaluar y contrastar el rendimiento de los modelos de estimación neuronal implementados. 	Julio a Diciembre 2013

Tabla 1: Plan de Trabajo.

5. Marco Teórico

5.1. Inteligencia Computacional

Inteligencia Computacional (IC) es una rama de la inteligencia artificial centrada en el estudio de mecanismos adaptativos para permitir el comportamiento inteligente de sistemas complejos y cambiantes. Se presenta como una alternativa a la GOFAI ("Good Old-Fashioned Artificial Intelligence"), tratando de no confiar en algoritmos heurísticos tan habituales en la Inteligencia Artificial más tradicional. Dentro de la Inteligencia Computacional podemos encontrar técnicas como las Redes Neuronales, Computación Evolutiva, Swarm Intelligence, Sistemas Inmunes Artificiales o Sistemas difusos. También se relaciona con técnicas como los Fractales, Teoría del Caos, Wavelets, Autómata celular, etc. La Inteligencia Computacional combina elementos de aprendizaje, adaptación, evolución y Lógica difusa para crear programas que son, en cierta manera, inteligentes. La investigación en Inteligencia Computacional no rechaza los métodos estadísticos, pero muy a menudo aporta una vista complementaria. Las Redes Neuronales son una rama de la inteligencia computacional muy relacionada con el aprendizaje automático. [14]

5.2. Redes Neuronales Artificiales

Una de las cosas que diferencia al ser humano del resto de los seres vivos es su capacidad de aprender y su inteligencia, las que le permite crear cosas de tal complejidad que están lejos del alcance de cualquier otra especie que habite este planeta. Es así que con su inteligencia el ser humano ha desarrollado a lo

largo de su historia objetos que facilitan su vida, por lo que desde la invención de la rueda, el ser humano no ha parado y ha llegado a crear inventos con la capacidad de realizar diversas tareas, ya sea de forma mecánica y/o lógica. Pero hay un límite en las invenciones mencionadas, estas realizan solo tareas para las que fueron creadas sin la capacidad de aprender, siguiendo alguna secuencia de pasos ya establecidos, de modo que surge la siguiente pregunta, ¿Cómo superar ese límite?, es aquí que el ser humano mira hacia él mismo y encuentra en su naturaleza la respuesta. La manera de superar el límite planteado es su capacidad de aprender y actuar acorde a lo aprendido. Así entonces la solución a la limitante anterior consiste construir sistemas inteligentes, este objetivo es precisamente el objetivo de la inteligencia artificial. En la inteligencia artificial se pueden apreciar dos áreas significativas:

- Inteligencia artificial simbólica.
- Inteligencia artificial sub simbólica.

La primera de ellas se encarga de construir sistemas que contenga características que se pueden determinar cómo inteligentes, para lo cual es preciso definir el problema y diseñar un sistema capaz de resolverlo siguiendo esquemas prefijados por la disciplina. Un ejemplo son los sistemas expertos. En la segunda de ellas no se hace un diseño de alto nivel de sistemas que resuelvan los problemas utilizando las técnicas de la disciplina, sino que se usa un sistema genérico que se va adaptando y construyéndose para formar por si mismo un sistema capaz de resolver el problema. Es en la inteligencia artificial sub simbólica donde se estudian los mecanismos del sistema nervioso, cerebro, sus características,

estructura y funcionamiento con el fin de diseñar programas que recojan esto y se adapten generando sistemas que puedan resolver problemas. Es en este campo en donde se encuentran las redes neuronales artificiales. Las redes neuronales artificiales tienen como objetivo emular el comportamiento de una red neural biológica, de manera de poder resolver problemas, para lo cual toma ciertos elementos de la red neuronal biológica como lo son las neuronas y crea un elemento análogo a este llamado neurona artificial, es por ello que para conocer el funcionamiento de una red neuronal artificial, es preciso adentrarse en los fundamentos biológicos que son la base para la construcción de una red neuronal artificial. Teniendo claro lo anterior es que posteriormente se ahondará más en profundidad sobre los fundamentos biológicos de las redes neuronales así como los principales elementos que son considerados a la hora de emularle en una red neuronal artificial. [9]

5.2.1. Reseña Histórica

La historia de las redes neuronales se remonta hacia el año 1936 en donde Alan Turing fue el primero en estudiar el cerebro como una forma de ver el mundo de la computación. Sin embargo, los primeros teóricos que concibieron los fundamentos de la computación neuronal fueron: el neurofisiólogo Warren McCulloch y el matemático Walter Pitts quienes, en 1943, describen una lógica de cálculo de redes neuronales que une estudios de neurofisiología y lógica matemática. Su modelo formal de una neurona sigue la ley "todo o nada". Con un suficiente número de estas simples unidades, correctas conexiones sinápticas y operando de manera síncrona, demostraron que una neurona constituida de

esta manera puede, en principio, calcular cualquier función computable.

En 1949 Donald Hebb, fue el primero en explicar los procesos del aprendizaje desde un punto de vista psicológico, desarrollando una regla de cómo el aprendizaje ocurría, planteaba que el aprendizaje se daba cuando ocurrían ciertos cambios en una neurona activando éstas. También propone que la conectividad del cerebro cambia continuamente conforme un organismo aprende cosas nuevas, creándose asociaciones neuronales con estos cambios. Los trabajos de Hebb formaron las bases de la Teoría de las Redes Neuronales.

En 1950 Karl Lashley resume su investigación de 30 años y destaca que el proceso de aprendizaje es un proceso distribuido y no local a una determinada área del cerebro.

Haibt y Duda en 1956 realizan una de las primeras simulaciones computacionales para probar una bien formulada teoría neuronal basándose en el postulado de aprendizaje de Hebb. Este mismo año, Uttley, demuestra que una red neuronal con sinapsis modificables puede aprender a clasificar un simple conjunto de patrones binarios en sus respectivas clases.

Quince años luego de la publicación de McCulloch y Pitts un nuevo acercamiento al problema de reconocimiento de patrones fue introducido por Rosenblatt (1958) en su trabajo sobre el perceptrón. Éste constaba de 2 niveles y los pesos se ajustaban en proporción al error entre las salidas deseadas y obtenidas. Dos años después Widrow y Hoff introdujeron el algoritmo "least mean-squarez lo usaron para formular el ADALINE (Adaptative Linear Element).Éste último, se diferencia del perceptrón en el proceso de entrenamiento.

En 1969 se produjo la "muerte abrupta" de las redes neuronales cuando Minsky

y Papera probaron matemáticamente que el Perceptrón no era capaz de resolver problemas tan fáciles como el aprendizaje de una función no lineal. Esto demuestra una gran debilidad, dado que las funciones no-lineales son ampliamente utilizadas en computación y en problemas del mundo real.

En 1982 Hopfield publica un trabajo clave para el resurgimiento de las redes neuronales. En él, desarrolla la idea del uso de una función de energía para comprender la dinámica de una red neuronal recurrente con uniones sinápticas simétricas. El principal uso de estas redes ha sido como memorias y como instrumento para resolver problemas de optimización como el problema del viajante. En 1986 Rumelhart, Hinton y Williams, desarrollan el algoritmo de aprendizaje de retro propagación (backpropagation) para redes neuronales multicapa, dando una serie de ejemplos en los cuales se demuestra el potencial del método desarrollado.

A partir de ese año, el número de trabajos sobre redes neuronales ha aumentado exponencialmente apareciendo un gran número de aportaciones tanto a los métodos de aprendizaje como a las arquitecturas y aplicaciones de las redes neuronales. Se podría destacar entre todas estas aportaciones los trabajos de Broomhead y Lowe y el de Poggio y Girosi sobre el diseño de redes neuronales en capas usan RBF (Radial Basis Function).

En la actualidad, son numerosos los trabajos que se realizan y publican cada año, las aplicaciones nuevas que surgen (sobre todo en el área de control) y las empresas que lanzan al mercado productos nuevos, tanto hardware como software (sobre todo para simulación).

5.2.2. Fundamentos Biológicos

Las redes neuronales artificiales se basan en el sistema de comunicación neuronal, el cual está formado por el sistema hormonal y por el sistema nervioso, estos sistemas se conectan con diversos órganos, lo que permitiendo obtener información del medio, procesar y almacenar información para posteriormente enviarla como órdenes a los diferentes órganos efectores. Este sistema está formado de tres partes:

- Receptores: estos se ubican en las células sensoriales y como su nombre lo indica son los encargados de recibir los estímulos del medio (externo o proveniente del interior del organismo).
- Sistema nervioso: encargado de recibir la información obtenida por los receptores, procesarla, almacenarla y enviar información a los órganos efectores como a otras partes del sistema nervioso.
- Órganos efectores: son los que reciben la información enviada por el sistema nervioso para así realizar la acción.

5.2.3. La Neurona Biológica

La célula nerviosa llamada neurona es la unidad estructural y funcional más esencial del sistema de comunicación neuronal. Cada una de ellas debe desempeñar cuatro funciones:

- Recibir información del entorno, ya sea interno o externo, o de otras neuronas.
- Integra la información recibida y producir una señal apropiada.

- Conducir la señal a su terminal nerviosa.
- Transmitir la señal a otras células nerviosas, glándulas o músculos.

Aun cuando existe variación en las neuronas, se pueden definir en todas ellas cuatro regiones bien definidas que realizan las funciones ya mencionadas. Dichas regiones son las dendritas, el cuerpo celular, el axón y las terminales sinápticas (figura 1).

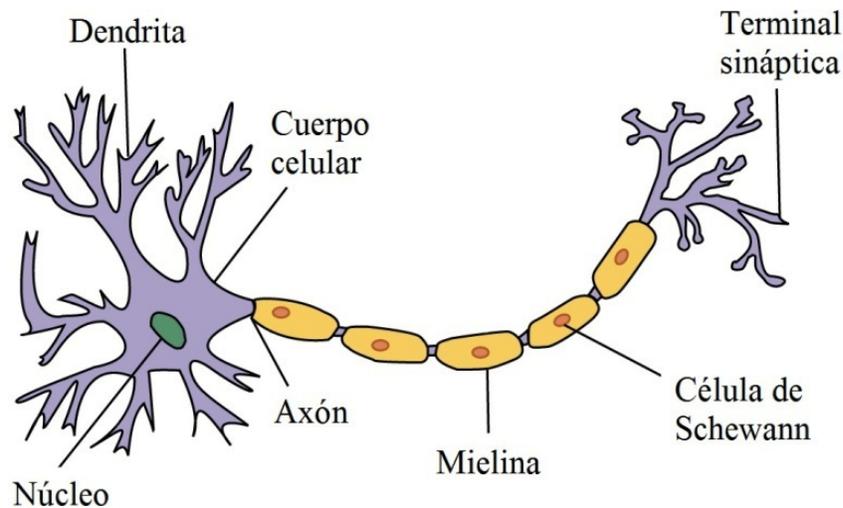


Figura 1: Neurona biológica

Dendritas: son prolongaciones del cuerpo de la célula nerviosa, ellas responden a señales de otras neuronas o del entorno exterior, encargándose de recibir información. En las neuronas del cerebro y la medula espinal, responden a los neurotransmisores químicos liberados por otras neuronas. Estas dendritas tienen en su membrana receptores proteicos que se unen a neurotransmisores específicos y, como resultado producen señales eléctricas. Las dendritas de las neuronas sensoriales tienen ciertas adaptaciones en su membrana permitiéndoles producir señales eléctricas en respuesta a estímulos específicos del entorno, como presión,

luz, calor o moléculas odoríferas.

Cuerpo celular: encargado de cumplir la segunda función, que es la de integrar la información recibida y producir una señal apropiada. A él viajan las señales recibidas por las dendritas, las que pueden ser positivas o negativas, las que se suman. Si la suma tiene una magnitud positiva suficiente, la neurona producirá un potencial de acción, el cual es la señal eléctrica de salida de la neurona, además en su interior alberga al núcleo, el cual se encarga de realizar actividades metabólicas de la neurona.

Axón: Son líneas de distribución que llevan potenciales de acción desde el cuerpo celular a las terminales sinápticas. La mielina es un aislante que cubre el axón, funcionando como un aislante que acelera la conducción de la señal eléctrica.

Terminales sinápticas: son engrosamientos en los extremos ramificados de los axones, son las terminales las encargadas de transmitir la señal a otras células. El punto en que las terminales sinápticas se comunican con otras células se llama sinapsis.

Una neurona inactiva tiene un potencia eléctrico constante, este potencial es llamado potencial de reposo, siempre es negativo dentro de la célula, variando entre -40 y -90 minivolts. Si la célula recibe un estímulo, su potencial se puede hacer más o menos negativo. Si esta magnitud negativa se reduce lo suficiente

como para alcanzar un nivel llamado umbral, ocurrirá un potencial de acción, en el cual el potencial de la neurona se elevara rápidamente hasta cerca de los 50 minivolts dentro de la célula. El potencial de acción dura unos pocos milisegundos, volviendo posteriormente la célula a su potencial normal. La carga positiva del potencial de acción fluye por el axón hasta las terminales sinápticas, donde la señal se comunica a otra célula a través de una sinapsis.

Cuando un potencial de acción llega a una terminal sináptica, el interior de la terminal adquiere una carga positiva. Esta carga tiene como efecto la liberación de neurotransmisores químicos al espacio entre las células, estos neurotransmisores se difunden rápidamente al otro lado de ese espacio y se unen a receptores en la membrana de la neurona post sináptica antes de alejarse por difusión y ser reabsorbidos por la neurona pre sináptica, así se transmite la señal entre neuronas.

Si la suma que se obtiene de los potenciales eléctricos mediante las dendritas alcanza un valor entre los 10 a 15 mini volts se alcanzara el umbral.

En el cerebro existen alrededor de 100000 millones de neuronas, donde cada una de ellas se conecta con alrededor de 10000 neuronas, es decir que cada actividad neuronal afecta a otras 10000 neuronas, con lo cual se forma una red de un gran tamaño. En la actualidad un dispositivo así está lejos del alcance. Los intentos más cercanos son de un millón de procesadores, donde cada procesador se conecta solamente con ocho adyacentes. Los sistemas artificiales actualmente

existentes se ven bajo esta limitante.

5.2.4. La Neurona Artificial

La neurona artificial es una emulación a menor escala de la neurona biológica. Posee un estado interno, el cual se conoce como nivel de activación, este estado puede cambiar según las señales que reciba, las que pueden provenir del medio o de otras neuronas artificiales.

El conjunto de estados que puede poseer la neurona se denomina S , el cual puede tener valores binarios, naturales o incluso valores en un intervalo continuo.

La neurona tiene un nivel de activación, el cual está determinado por las entradas y los pesos sinápticos que esta posea, si bien las entradas son externas a ella, los pesos pueden cambiar para así adaptar el nivel de activación. Este cambio va a depender de las entradas que recibe la neurona mediante una función de activación, estas entradas pueden venir de otras neuronas o del exterior.

Como se desea calcular el nivel de activación, es preciso en primer lugar calcular la entrada total de la célula E_i , este valor se calcula mediante la ecuación 2, en donde se suma cada entrada ($X = [x_1, x_2, \dots, x_j, \dots, x_n]$, representando las señales de la sinopsis en una neurona biológica) multiplicada por un peso sináptico ($W = [w_{i1}, w_{i2}, \dots, w_{ij}, \dots, w_{in}]$, representando la fuerza de una conexión sináptica) y se le resta un umbral θ .

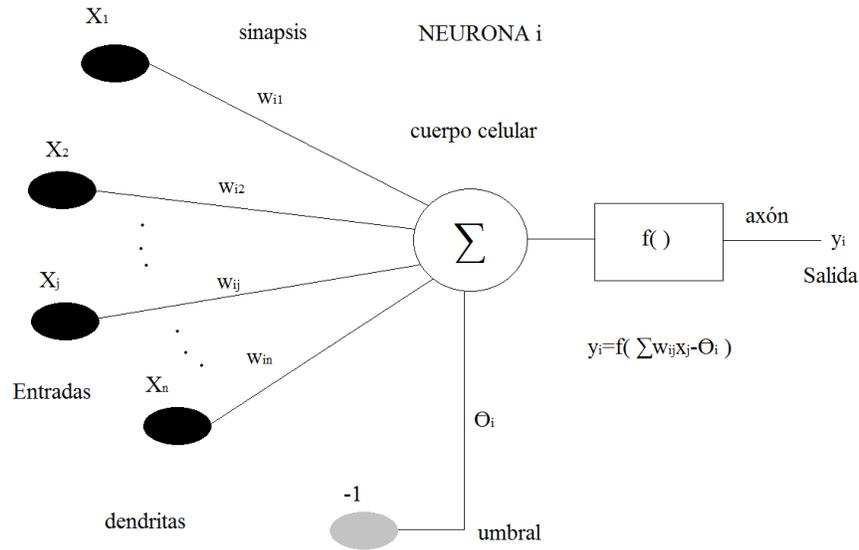


Figura 2: Neurona artificial

$$E = (X^T \times W) - \theta = x_1 \times w_1 + \dots + x_n \times w_n - \theta = \sum_{i=1}^N w_i \times x_i - \theta \quad (2)$$

La entrada total de la célula, E , es procesada por la función de activación $f()$, produciendo una señal de salida S . Es así que dependiendo de la función $f()$ es que se tendrán diferentes modelos de autómatas.

$$S = f(E) \quad (3)$$

5.2.5. Clasificación de Redes Neuronales

Las neuronas artificiales se unen para formar una estructura neuronal o modelo conexionista, los que pueden ser clasificados según [15]:

El número de capas:

- Redes neuronales mono capa: esta clase de redes son las más simples, si bien estas redes tienen dos capas, la primera capa se encarga de recoger las entradas sin realizar cálculos sobre ellas por lo que no se considera como una capa para efectos de cálculo, es en la segunda capa (capa de salida) donde se realizan los todos los cálculos.
- Redes neuronales multicapa: es una generalización de la anterior, en este tipo de redes existen un número de capas entre la capa de entrada y la capa de salida llamadas capas ocultas.

El tipo de conexiones:

- Redes neuronales no recurrentes: Esta estructura no posee memoria pues las señales se propagan solamente hacia adelante, sin existir la posibilidad de realimentación. Estas redes también se conocen como redes de propagación hacia adelante, en inglés Feed-Forward.
- Redes neuronales recurrentes: esta estructura es adecuada para el estudio de la dinámica de sistemas no lineales. En ella se permite la realimentación, ya sea entre neuronas de la misma capa, de diferentes capas o de la neurona misma. Estas redes también se conocen como redes de retro propagación, en inglés Feed-Back.

El grado de conexión:

- Redes neuronales totalmente conectadas: Todas las neuronas de una capa se encuentran totalmente conectadas con las de la capa siguiente o anterior.
- Redes parcialmente conectadas: En este caso no se da la conexión completa entre las neuronas de diferentes capas.

5.2.6. Estructura de una Red

Anteriormente se entró en detalle respecto a la neurona artificial, pero este elemento por si solo no es mucho lo que puede hacer, mas al conectarse estas unidades forman un elemento más complejo llamado red neuronal artificial. Estas neuronas se organizan en capas, en donde cada capa cumple una función, es así que tenemos tres tipos de capas dentro de una red neuronal, la capa de entrada, la capa oculta y la capa de salida. Como se puede observar en la figura 3, la red parte con la capa de entrada, esta capa cumple la función de recibir los datos que vienen del exterior de la red neuronal propagando los datos a la siguiente capa de la red sin modificar dichos datos, por lo que no se considera al momento de contabilizar el total de capas de la red neuronal, después viene la capa oculta, la cual se encuentra entre la capa de entrada y la capa de salida de la red, esta capa puede albergar más de una capa oculta, esto puede favorecer a resolver problemas más complejos, pero aumentará la complejidad de la red, al final de la red neuronal se encuentra la capa de salida, la cual arroja los resultados de salida de la red neuronal.

Como se aprecia en la figura 3 se representa una red neuronal de tres capas, estas capas se conectan mediante conexiones, las cuales tendrán pesos sinápti-

cos. Para conocer cuántos pesos sinápticos posee la red basta con multiplicar la cantidad de neuronas de cada capa por la cantidad de neuronas de la siguiente capa, por ejemplo si tomamos la figura 3 obtendremos que la red tiene $2 \times 3 + 3 \times 2 = 12$ pesos sinápticos.

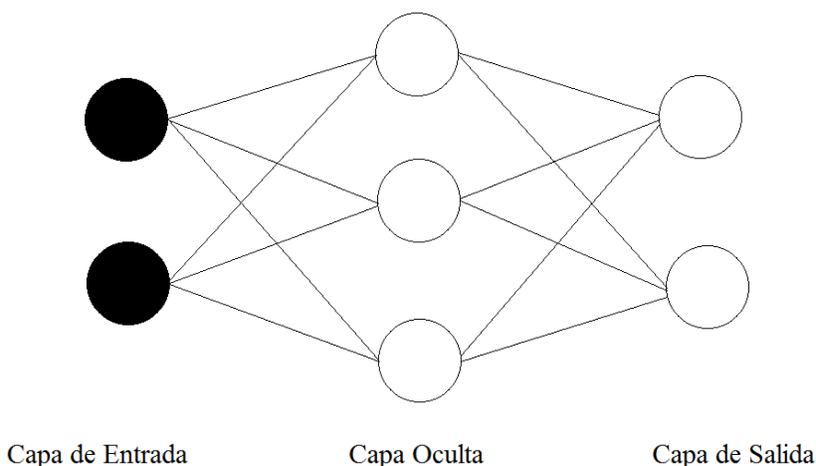


Figura 3: Red neuronal artificial de tres capas

Como se pudo ver, las neuronas artificiales al conectarse ofrecen un potente elemento, la red neuronal artificial, estas redes se caracterizan principalmente por [6]:

- La naturaleza de lo que emulan, tienen una inclinación a la adquisición de conocimiento por medio de las experiencias, este conocimiento es almacenado al igual que en el cerebro en el peso relativo de las conexiones inter neuronales.
- Poseer gran plasticidad y adaptabilidad, teniendo la capacidad de cambiar dinámicamente junto con el medio.

- Tener tolerancia a fallas, del mismo modo como ocurre con los sistemas biológicos, pueden sufrir daños considerables y aun así continuar teniendo un buen comportamiento.
- Tener un comportamiento no-lineal, esto les permite obtener y procesar la información proveniente de fenómenos no lineales.

Es así como se estructura a grandes rasgos una red neuronal artificial, si bien puede parecer simple respecto a la estructura neuronal biológica, este tipo de emulación ha logrado buenos resultados en diversas áreas.

5.2.7. Aprendizaje

El aprendizaje o entrenamiento es la fase más importante dentro de las redes neuronales artificiales, esto debido a que el tipo de problemas que podrá resolver la red estarán determinados por el esquema de aprendizaje ocupado. Esto se debe a que la red neuronal artificial es una emulación de la red neuronal biológica, y como tal su aprendizaje estará basado en los ejemplos con los que aprenda. Desde este matiz se puede decir que el conjunto de ejemplos con los que aprenda una red neuronal debe ser significativo y representativo. Significativo en cuanto a que la cantidad de ejemplos con los que se entrene la red debe ser lo suficientemente amplio como para que la red pueda adaptar sus pesos de manera eficaz, de lo contrario los resultados no serán satisfactorios. Representativo, en el sentido que los componentes del conjunto de aprendizaje deben ser diversos, de tal manera que la red no se especialice en un tipo de ejemplo por sobre el resto, puesto que implicaría que su aplicación no sea general.

Con el aprendizaje se pretende poder resolver problemas de manera eficaz, esto mediante la calibración de los pesos de todas las conexiones de la red.

Para calibrar los pesos se van introduciendo los ejemplos del conjunto de aprendizaje, con cada ejemplo los pesos se irán adaptando conforme el esquema de aprendizaje utilizado, si después de ingresados todos los ejemplos la red aun no converge conforme cierto criterio, entonces se vuelve a entrenar con el conjunto de ejemplos. La modificación de los pesos puede hacerse durante o después de ingresados todos los ejemplos.

El periodo de aprendizaje puede ser determinado según [12]:

- Un número fijo de ciclos: después de un cierto número fijo de ciclos se acepta la red resultante.
- El error sea inferior a una cantidad preestablecida. De ser el caso, primero se debe definir una función de error a nivel de patrón individual y a nivel de la totalidad del conjunto de entrenamiento. Este valor se decide a priori. Este criterio debe ser acompañado de algún otro como por ejemplo el anterior, esto en caso de que la red no converja (señal de que la red no ha alcanzado una solución y por ende requerirá modificaciones).
- La modificación de los pesos sea irrelevante. Básicamente es cuando la variación de los pesos no es significativa o no se produce variación de los pesos en ninguna conexión.

Los tipos de aprendizaje que se pueden dar son: aprendizaje supervisado, no

supervisado, por refuerzo e híbrido.

5.2.7.1. Aprendizaje Supervisado

El aprendizaje supervisado como su nombre lo sugiere, es un aprendizaje en el cual existe un control de las salidas, este tipo de control se puede hacer mediante tres formas:

- Aprendizaje por corrección: en el cual los datos de aprendizaje tienen dos tipos de atributos, los datos propiamente tal y datos relacionados a la solución del problema. La forma más habitual de modificar los pesos de las conexiones se puede observar en la figura 4. Cuando se introduce un ejemplo este es procesado, la salida obtenida es comparada con la salida esperada. La diferencia entre la salida esperada y la obtenida influirá en la modificación de los pesos, es decir, si la diferencia es leve los pesos variarán poco, de lo contrario la variación en los pesos será mayor.

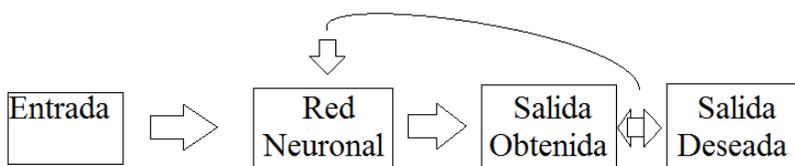


Figura 4: Aprendizaje supervisado

- Aprendizaje por refuerzo: en este caso solo se conoce si la salida de la red es o no correcta, en otras palabras la información es de tipo booleana (verdadero o falso)
- Aprendizaje estocástico: en este caso se hacen cambios aleatorios en los

valores de los pesos de las conexiones de la red y se evalúa su efecto según el objetivo deseado y distribuciones de probabilidad. Se suele hacer una analogía con termodinámica, pues se considera un estado de energía E el que simboliza la estabilidad. Si después de realizar los cambios el estado de energía es menor, se aceptara el cambio, en caso contrario se aceptaría solo dependiendo de la distribución de probabilidad.

5.2.7.2. Aprendizaje No Supervisado

Este tipo de aprendizaje a diferencia del anterior no posee los datos relacionados con la solución del problema, por lo que no hay datos que permitan guiar el proceso de aprendizaje. Por lo cual los pesos de la red se modificarán a partir de la información interna, entonces la red tratará de determinar características de los datos del conjunto de entrenamiento.

Por lo general suelen haber dos tipos de aprendizaje no supervisado:

- Aprendizaje Hebbiano: en él los pesos de las conexiones se ajustan según la correlación que halla entre las neuronas. Si las dos neuronas son activas (positivas) se reforzará la conexión, pero si una es activa y la otra pasiva (negativa) se debilita la conexión.
- Aprendizaje Competitivo o cooperativo: En este caso las neuronas cooperan y compiten entre ellas para realizar una tarea. Al presentar información de entrada, solo un grupo o una neurona de salida se activará, anulando a las demás.

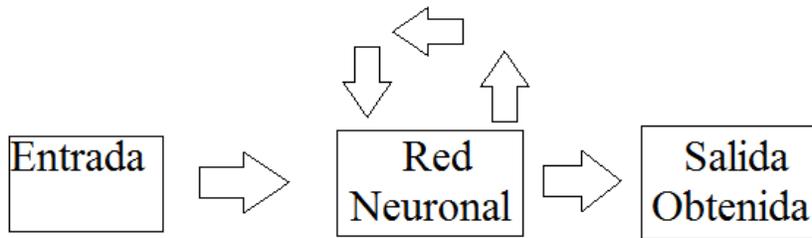


Figura 5: Aprendizaje no supervisado

5.2.7.3. Aprendizaje Híbrido

Como su nombre lo indica es una mezcla de aprendizajes, esta se realiza utilizando el aprendizaje supervisado y el aprendizaje no supervisado. Este aprendizaje puede ser empleado en redes neuronales artificiales de base radial [12], para ello se utilizan dos fases, en la primera (fase no supervisada) se determinan los centros y amplitudes de las neuronas en la capa oculta, para luego en la segunda fase (fase supervisada) determinar los pesos y umbrales de la capa de salida.

5.3. Optimización por Enjambre de Partículas

La optimización por enjambre de partícula, más conocida en la literatura científica como *Particle Swarm Optimization* (PSO), nace, al igual que otras técnicas estocásticas del calculo evolutivo, en un intento por imitar y mimetizar el comportamiento de los procesos naturales. Siendo uno de los métodos más utilizados en la inteligencia computacional, busca imitar comportamientos sociales de un colectivo a partir de la interacción entre sus individuos y de éstos

con su entorno.



Figura 6: Enjambres en la Naturaleza

Fue desarrollado por Kennedy y Eberhart en 1995 [7], basando su inspiración en el comportamiento de sistemas biológicos sociales como lo son las bandadas de pájaros, cardúmenes de peces e insectos entre otros (figura 6). Ya que trata de simular el comportamiento de estos sistemas biológicos sociales en situaciones tales como búsqueda de alimento o de lugar de cobijo. Para ello se considera que cada individuo tiene un comportamiento determinado tanto por su búsqueda individual como por la búsqueda grupal. De este modo se logra llegar a una analogía en la cual se tiene un enjambre de partículas que simulan a los individuos en un sistema biológico social, estas partículas se desplazan en un espacio en el cual buscan soluciones.

5.3.1. Algoritmo PSO Modelo Tradicional

Definición: PSO es un algoritmo evolutivo, iterativo y estocástico, que se basa en buscar soluciones en un espacio vectorial N dimensional desplazando

a un conjunto de partículas sobre el mismo, siendo cada partícula una posible solución.

Este algoritmo parte de una población(o enjambre) para iniciar un proceso de búsqueda, donde cada partícula esta agrupada dentro una población con posiciones y velocidades aleatorias. Una partícula es considerada como un punto en el espacio de búsqueda D-dimensional, donde cada dimensión representa una de las incógnitas del problema. La nomenclatura para representar el enjambre está dada por $X = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{id})$, donde $i = 1, 2, \dots, n$ corresponde a la i-ésima partícula y $n = 1, 2, \dots, D$, corresponde a sus dimensiones.

Ya teniendo inicializadas las partículas comienza el proceso iterativo. Por cada iteración, cada partícula actualizara el valor de su mejor posición encontrada, denominada mejor personal p_{best} . Para representar al p_{best} de una i-ésima partícula se tiene la siguiente expresión $P_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{id})$. Además de actualizar él p_{best} se debe preocupar de actualizar el valor de la mejor posición encontrada en el enjambre, denominado mejor global p_{gbest} . Estos valores se utilizan para ajustar la velocidad con que se mueven una partícula i-ésima en cada una de las dimensiones. Además, con el valor de la velocidad, la partícula podrá calcular su nueva posición para la siguiente iteración. La influencia que tiene la mejor posición personal sobre la velocidad de una partícula se conoce como el factor o componente de cognición, y la influencia de la mejor posición global o del enjambre se conoce como componente social [1].

Las ecuaciones que determinan la posición y la velocidad de cada partícula dentro del enjambre está dada por:

$$v_{id}(t+1) = w \cdot v_{id}(t) + c_1 r_1 (p_{id}(t) - x_{id}(t)) + c_2 r_2 (p_{gd}(t) - x_{id}(t)) \quad (4)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (5)$$

Los parámetros c_1 y c_2 corresponde a las variables que controlan la influencia de las componentes cognitiva y social. Los valores de r_1 y r_2 corresponden a números aleatorios independientes pertenecientes al intervalo $[0, 1]$. Sin embargo hasta ahora no se ha hablado del el parámetro w . En un trabajo previo [8], se realizó una modificación al algoritmo agregando un componente denominado factor de inercia w , el cual se utiliza para balancear la velocidad de la partícula respecto a la búsqueda global. Para prevenir que las velocidades de las partículas incrementen infinitamente, estas se acotan a un rango determinado por el valor V_{max} , el cual delimita el rango de velocidad al intervalo $[-V_{max}, V_{max}]$.

A continuación se mostrara los pasos a seguir en el algoritmo Optimización por Enjambre de Partículas (PSO) tradicional:

1. Se inicia con una población de partículas de tamaño n , con posiciones y velocidades aleatorias en un espacio D-dimensional.

2. Para cada partícula, evaluar su función fitness de acuerdo a algún criterio de optimización (función de costo, mínimos cuadrados, error cuadrático medio).
3. Se compara el valor de la función fitness de la partícula con el fitness del mejor personal (p_{best}). Si el valor del fitness actual resulta ser mejor, entonces se actualiza p_{best} al valor de la posición actual x_{id} .
4. Se compara el fitness actual de cada partícula con el mejor *fitness* encontrado por el enjambre. Si el *fitness* actual resulta ser mejor, entonces se actualiza la mejor posición global (p_{gbest}) al valor de la posición actual x_{id} .
5. Modificar las posiciones y velocidades de las partículas de acuerdo a las ecuaciones 4 y 5, verificando que las partículas no sobrepasen la velocidad máxima definida el intervalo generado por V_{max} .
6. Si el criterio de parada no se ha cumplido, volver al paso 2.

Las condiciones de términos más recurridas en la literatura para finalizar el proceso de optimización son:

- Número de iteraciones sin mejora: El proceso termina después de alcanzar un número fijo de iteraciones sin presentar mejoras en la solución.
- Número máximo de iteraciones sin mejora: El proceso termina después de realizar un número fijo de iteraciones preestablecidas.
- Fitness logrado: El proceso termina cuando llega a un valor menor o igual al fitness establecido.

El éxito del algoritmo depende en la medida que se ajusten las posiciones de las partículas en el espacio de búsqueda, estos ajustes dependen de un número reducido de parámetros lo que se traduce en la gran potencia de PSO. Sin embargo muchísimas variaciones de los parámetros se podrían hacer. Es por ello que se a continuación se presenta un apartado para revisar las posibles modificaciones que se aconsejan en la literatura para ajustar los parámetros y obtener soluciones prometedoras.

5.3.2. Ajuste de los Parámetros de PSO

El ajuste de los parámetros condiciona el rendimiento posterior del algoritmo de optimización. Existen múltiples estudios [1] que intentan generalizar la selección de estos parámetros, sin embargo dicha selección está directamente relacionada a la naturaleza del problema.

Retomando la ecuación 4, se podrían variar sus parámetros de manera tal que afecte la búsqueda en el espacio de soluciones. A continuación se examinarán los parámetros que afectan al algoritmo.

$$v_{in}(t+1) = \underbrace{w \cdot v_{in}(t)}_{\text{Inercia}} + \underbrace{c_1 r_1 (p_{in}(t) - x_{in}(t))}_{\text{Influencia Personal}} + \underbrace{c_2 r_2 (p_{gn}(t) - x_{in}(t))}_{\text{Influencia Social}} \quad (6)$$

- **Velocidad:** Para acotar la velocidad de la partícula se especifica un valor máximo, V_{max} , que restringe la velocidad en cada dimensión al intervalo

$[-V_{max}, V_{max}]$. Si el valor de V_{max} es extremadamente pequeño las partículas explorarían el espacio de soluciones muy lentamente y podrían quedar atrapadas en un óptimo local, este fenómeno se conoce como convergencia prematura del algoritmo. No incluir el parámetro de control V_{max} , implicaría que el enjambre no tienda a converger hacia un punto, produciéndose el fenómeno conocido como explosión del PSO, consistente en un comportamiento oscilatorio y creciente de la posición de las partículas, provocando la ineficiencia del PSO como algoritmo de optimización [8]. El valor de V_{max} a tomar, dependerá única y exclusivamente de la naturaleza del problema a optimizar. Con el objetivo de reducir el efecto de V_{max} y perfeccionar el control del alcance de la búsqueda sobre el espacio de soluciones se incorporó el concepto de peso inercial [1].

- **Peso de inercia w :** Es considerado un parámetro crítico para la convergencia de PSO. Se utiliza para controlar el impacto de la historia previa de las velocidades actuales. En consecuencia, el parámetro w regula la compensación entre las capacidades de exploración global y locales del enjambre. Un peso inercial grande facilita la exploración global (buscando nuevas áreas). Si es un valor pequeño facilita la exploración local. Algunos resultados experimentales han indicado que es mejor asignar un valor w grande al principio, para promover la exploración global del espacio de búsqueda, y gradualmente disminuirlo para conseguir soluciones más refinadas. En [1] se menciona que una buena opción para el parámetro w sería asignarle un valor inicial alrededor de 1.0 y luego, ir disminuyéndolo hasta llegar cercano a 0. Otra alternativa, sería usar alguna aproximación adaptable (un contro-

lador difuso por ejemplo), en el cual este parámetro puede ser sintonizado según las características del problema a considerar.

- **Parámetros c_1 y c_2 :** No son críticos para la convergencia, sin embargo, si se calibran adecuadamente podrían acelerar la convergencia del algoritmo y aliviarlo de posibles mínimos locales. El valor comúnmente usados para estos parámetros es igual a 2 ($c_1 = c_2 = 2$) [1], pero algunos resultados experimentales indican que se pueden asignar a ambos parámetros el valor de 1.49. Trabajos recientes han agregado la restricción que la suma de los parámetros no supere el valor 4, es decir, $c_1 + c_2 \leq 4$ [8].
- **Tamaño de la población** (o enjambre): Debe ser seleccionado con rigurosidad ya que poblaciones excesivamente grandes exploran minuciosamente el espacio de búsqueda, pero con un elevado coste computacional. Los estudios paramétricos revelan que una población de alrededor de 30 partículas es suficiente para múltiples problemas, y típicamente se utilizan poblaciones que oscilan entre 10 y 50 partículas, o entre 100 y 200 partículas para abordar problemas más complejos [1].

Lo ideal es que las partículas representen soluciones de diferentes zonas del espacio de búsqueda. Este concepto se conoce diversificación de la población, y se refiere a que las partículas sobrevuelen diferentes zonas del espacio de solución antes de que se comprometan con alguna zona. Una vez que convergen a una zona prometedora, se espera que sean capaces de mejorar la solución encontrada es decir, que logren encontrar y converger a un óptimo cada vez mejor [1].

Estos 2 modos se conocen en la literatura exploración y explotación [1]. Se define como exploración, a la búsqueda de la solución del problema en un espacio de búsqueda amplio, es decir, que las partículas vuelen por diferentes zonas del espacio de búsqueda privilegiando la diversificación. Explotación se entenderá como a la búsqueda de la solución del problema encontrando sub-espacio dentro del espacio de búsqueda, en donde el enjambre trate de tomar todos los valores posibles que comprendan a sub-espacio. .

Lo ideal es ajustar los parámetros para que el algoritmo funcione en modo exploración inicialmente, y que luego cambie al modo explotación. Debido a esto, en la actualidad se han propuesto una serie de técnicas, en donde los parámetros son autoajustados durante la ejecución del algoritmo, cambiando de modo exploración a explotación y de explotación a exploración, según como sea el rendimiento del algoritmo .

En un trabajo previo [8], se da cuenta que valores altos para el coeficiente de inercia, factores de cognición y V_{max} favorecen el modo exploración del PSO, mientras que valores bajos para estos parámetros favorecen el modo explotación, es decir:

$$\textit{Exploración} \Rightarrow 0,75 < w < 1, 2 < c_1, c_2 < 4 \quad (7)$$

$$\textit{Explotación} \Rightarrow 0,4 < w < 0,75, 0,1 < c_1, c_2 < 2 \quad (8)$$

Finalmente una de las decisiones claves es la elección de la función objetivo para evaluar la aptitud o fitness de una partícula. La elección de esta función será clave para alcanzar el éxito del algoritmo, debido a que es ella quien se encarga de guiar la búsqueda durante el proceso de optimización.

5.3.3. PSO con Búsqueda Local Adaptativa (LPSO)

En el PSO estándar, las partículas son en gran parte influenciadas por su mejor posición alcanzada en su recorrido y por la mejor posición del enjambre. Una vez que la mejor partícula, no presenta cambios en un óptimo local, el resto de las partículas convergerán rápidamente a la posición presentada por la mejor partícula, esto se puede lograr a través de búsquedas sobre la vecindad de la mejor partícula del enjambre en cada generación.

Se propone un nuevo operador de búsqueda local adaptativa, llamado también operador de mutación, el cual consiste en mutar en cada iteración la mejor partícula encontrada hasta ese momento, la cual variará respecto a las ecuaciones presentadas a continuación:

$$gBest_j(t + 1) = gBest_j(t) + [\beta_j(t) - \alpha_j(t)] * rand() \quad (9)$$

$$\alpha_j(t) = \min(x_{ij}(t)); \quad \beta_j(t) = \max(x_{ij}(t)) \quad (10)$$

$$i = 1, 2, \dots, ps; \quad j = 1, 2, \dots, n \quad (11)$$

Donde $gBest_j$ es el vector de la partícula que presenta la mejor solución del

enjambre, $\alpha_j(t)$ y $\beta_j(t)$ son los valores mínimos y máximo de la dimensión j en el espacio de búsqueda, mientras que $rand()$ es un número aleatorio entre 0 y 1, ps corresponde al tamaño de la población e t indica en número de iteraciones. El tamaño de este operador ($\alpha_j(t) - \beta_j(t)$) puede ser considerado como el diámetro del espacio de búsqueda de la población. Al principio de la evolución, el espacio de búsqueda inicial es grande, por ende, el valor del operador es demasiado alto, siendo esto beneficioso para una búsqueda global, acelerando la convergencia del algoritmo. Sin embargo, con el aumento de las iteraciones, la población poco a poco converge a la mejor posición encontrada hasta ese momento. Por lo que el espacio de búsqueda de la población disminuye, así como también, el valor para el operador de búsqueda. Con esta situación, se beneficiará una búsqueda local en el espacio de soluciones, lo que conlleva a explotar de mejor manera la vecindad de la mejor partícula del enjambre.

5.4. Análisis de los Accidentes de Tránsito

Por su parte, Mussone, Ferrari y Oneta en 1999 [10] utilizaron redes neuronales para analizar accidente ocurridos en intersecciones en Milán, Italia. Eligieron las técnicas de feed-forward (MLP) con Back Propagation (BP). El modelo contó con 10 nodos de entrada para ocho variables día o noche; flujos de tráfico en las intersecciones; número de puntos virtuales conflictivos; número de puntos reales conflictivos; tipo de intersección; tipo de accidente; estado de superficie de las carreteras y condiciones climáticas. Los nodos de salida fueron llamados "índice de accidente", los cuales fueron calculados como un ratio entre el número de accidentes para una cierta intersección y el número de accidentes para

la intersección más peligrosa. Los resultados mostraron que el mayor índice de accidentes ocurrió en intersecciones no señalizadas en la noche.

El año 2002 So Young Sohn y Sung Ho Lee [16] usaron red neuronal y árbol de decisión para mejorar la exactitud de clasificación individual. Esto se desarrollo para dos categorías de gravedad en accidentes de tránsito, lesiones corporales y daño de propiedad. Se usaron tres enfoques distintos (classifier fusión basado en el algoritmo de Dempster- Shafer, procedimiento Bayesiano y modelo logístico); fusión de datos basado en arcing y bagging. El ultimo enfoque fue clustering basado en el algoritmo k-means. Los resultados obtenidos se midieron en base a la precisión obteniéndose como resultados un 76,10 % en el método de Clustering (con árbol de decisión) seguido de 74.78 % en el método bagging (con árbol de decisión). Las redes neuronales dieron como resultado un 70.86 % de precisión.

Se realizo una clasificación de accidentes según la severidad del accidente, clasificando este en cinco clases (no lesión, posible lesión, lesiones no incapacitantes, lesiones incapacitantes y lesiones fatales) y se compararon las técnicas de árbol de decisión, red neuronal (usando un entrenamiento que combinaba Back Propagation y conjugate gradient descent), hibrid DT-ANN y SVM usando la función de base radial para el kernel. Como resultado se obtuvo que en todas las clases el árbol de decisión presento mejor precisión. Los resultados se presentan en el informe referenciado [4].

En el año 2008 Juan Zúñiga presentó un modelo de redes bayesianas para identificar los factores determinantes en los accidentes de tránsito que afecten a las personas. Los registros de los accidentes contienen principalmente la fecha en la que ocurrieron éstos y sus características (Causas, Estado Atmosférico, Personas Involucradas, Vehículos Involucrados, etc.). Dentro del informe se muestran los resultados obtenidos, los cuales brindan una variedad importante de causas que provocan los accidentes de tránsito, el detalle de los resultados obtenidos se muestran en el informe referenciado [11].

Se realizó un estudio [2] que comparó dos modelos analíticos (una modificación del modelo Smeed y el modelo de Andreassen adaptado) y una red neuronal artificial para estimar el número de muertos en accidentes de tránsito, para ellos se probaron utilizando los datos históricos de 20 años de tres grandes metrópolis de Turquía, los resultados mostraron que la red neuronal artificial predecía mejor los resultados al dar los menores errores.

El último informe de seguridad vial realizado por la organización mundial de la salud en el 2009 [13], sostuvo como conclusión que la tasa más alta de letalidad por cien mil habitantes correspondían a países de ingresos bajos y medios. Sin embargo, este estudio no utilizó herramientas de minería de datos sino que solo técnicas y herramientas estadísticas.

En Julio del 2010 Félix Castro presentó un modelo de clasificación de accidentes de tránsito [3], para lo cual usó un modelo LS-SVM basado en Optimización

por enjambre de partículas. Este modelo se basa en el hecho de seleccionar los parámetros que usara LS-SVM mediante un algoritmo de PSO, una vez obtenidos estos parámetros se vuelve a entrenar para obtener una buena clasificación. Para este trabajo se clasificó en lesionados e ilesos, dando como mejor resultado una exactitud del 70 %.

Redes Neuronales, Arboles de Decisión, Redes Bayesianas y Clustering son técnicas predominantes dentro de los estudios realizados referentes a los accidentes de tránsito, sin embargo, los resultados obtenidos no son los más óptimos al momento de decidir si un modelo es realmente efectivo. Es por ello que se busca mejorar las técnicas ya realizadas con el fin de obtener mejores resultados y poder contribuir con un modelo que permita ser un aporte al estudio de la seguridad vial.

6. Desarrollo del Modelo

Para el estudio se utilizarán los datos suministrados por el CONASET, estos datos se encuentran almacenados en plantillas Excel separados por año.

Los datos se obtienen mediante su registro en plantillas SIEC2 para cada accidente, siguiendo un instructivo para posteriormente ser ingresados en una base de datos de CONASET.

Los datos son separables en tres entidades, cada una con sus respectivos atributos para fines de registro:

1. Accidentes: identificador del accidente, fecha, hora de accidente, comuna, causas, tipo de accidente, estado atmosférico, condición calzada, tipo de calzada, estado calzada, urbano/rural, y la información que indica la cantidad de ilesos, leves, graves, menos graves y muertos.
2. Persona: identificador del accidente, sexo, calidad, edad, y resultado.
3. Vehículo: identificador del accidente, servicio, y el tipo de vehículo.

Como se indica en la siguiente figura 7, en cada accidente puede haber una o muchas personas involucradas y/o uno o muchos vehículos involucrados.

Para esta investigación se usaron los siguientes atributos propios de la entidad de accidentes.

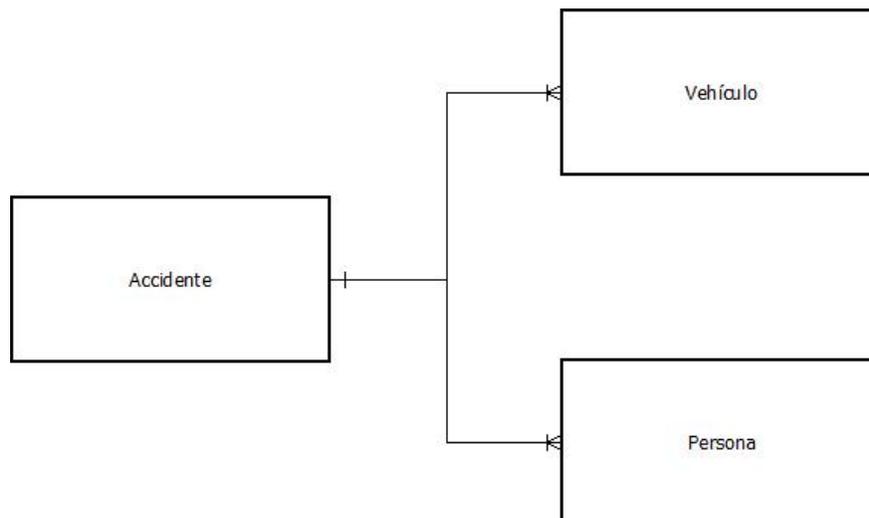


Figura 7: Entidades base de datos CONASET

Accidentes. Entidad correspondiente a los accidentes ocurridos en una determinada fecha y lugar geográfico. los atributos que se usaron son:

- Fecha: fecha del suceso.
- Comuna: comuna de la región en donde ocurrió el accidente.
- Muertos: cantidad de muertos involucrados en el accidente.
- Graves: cantidad de graves involucrados en el accidente.
- Menos Graves: cantidad de menos graves involucrados en el accidente.
- Leves: cantidad de leves involucrados en el accidente.

De los datos del accidente se trabajó con la fecha del accidente separando las cantidades de Muertos, Graves, Menos graves y Leves en semanas, los cuales representarán los datos de entrada. Para analizar los datos se separaron en dos grupos separados por las comunas. Valparaíso, Viña del Mar, Concón, Quilpué y Villa Alemana conforman el grupo denominado “Gran Valparaíso”, mientras

que el resto de las comunas pertenecientes a la V región conforman el grupo denominado “Otras Comunas”.

6.1. Graficos de representación de datos

En esta sección se motraran algunos graficos de analisis de los dato entre los años 2003 y 2012.

Cuando analizamos la cantidad de fallecidos (figura 8) podemos observar que el cada año va disminuyendo la cantidad a excepción de los años 2007 y 2011 donde se logra ver un aumento de la cantidad de fallecidos.

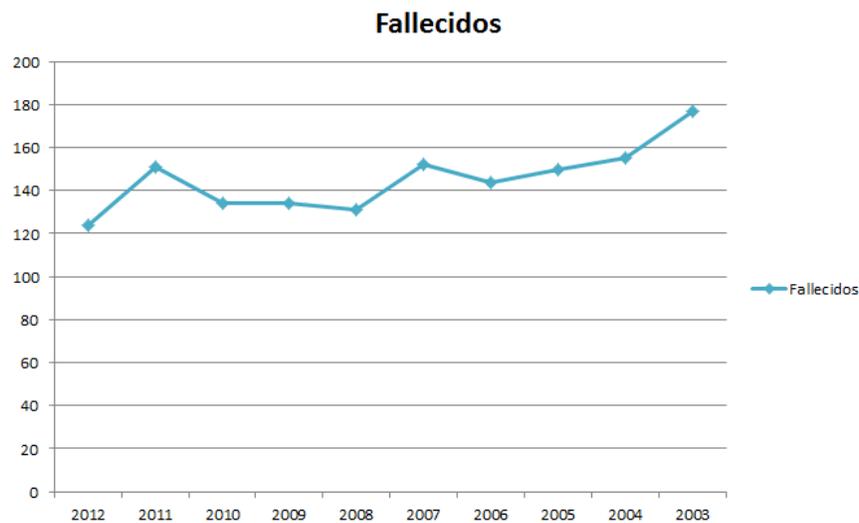


Figura 8: Cantidad de Fallecidos

La cantidad de lesionados como se puede observar en la figura 9, la cantidad de lesionados leves es mucho mayor a la cantidad de lesionados graves y menos graves, los cuales no varian mucho entre difrentes años. Con respecto a los lesionados leves va aumentando levemente a medida que va aumentando los años.

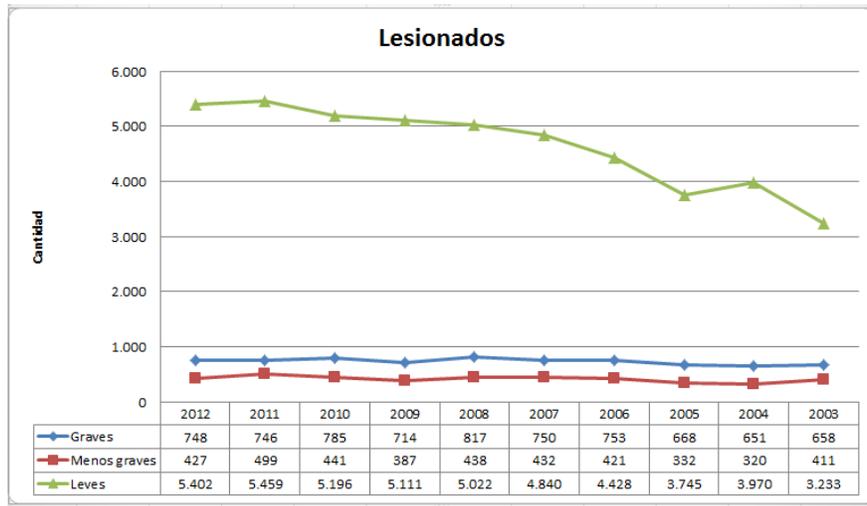


Figura 9: Cantidad de lesionados

6.2. Desarrollo del Modelo

Para llegar a la obtención de los modelos se ha de seguir una serie de pasos, estos se detallan a continuación:

1. **Tratamiento de datos:** cada accidente tiene una suma de todos los fallecidos, graves, menos graves y leves que están involucrados, junto a la fecha del accidente y causa que lo originó. Con esos datos se comenzó por sumar todos los fallecidos agrupando los datos por semana, en base a esos datos extraídos se realizó la ejecución de la red. La extracción de los datos fueron agrupados en 2 grupos de comunas, las denominadas “Gran Valparaíso” y las que no están en esa clasificación. Lo cual conlleva a realizar un análisis de 2 grupos por comunas.
2. **Selección de Algoritmo de entrenamiento:** se selecciona el algoritmo

de entrenamiento a utilizar que será PSO Modelo Tradicional y PSO con Búsqueda Local Adaptativa.

3. **Datos de entrada:** Para la red los datos de analisis son las cantidades de fallecidos y lesionados (que agrupan los graves, menos graves y leves). Estas cantidades son las entradas de la red que analizará.
4. **Elección de topologías:** los nodos de entrada de la red son el número de desfases que se quiere implementar en la red, se testeo con desfases desde 1 hasta 48, que para nuestra investigación equivale a un periodo de 48 semanas por cada año. Los nodos de la capa oculta se irán testeando desde 1 hasta 15 para obtener el mejor resultado. La capa de salida presenta un nodo de salida.
5. **Entrenamiento de la red:** la red se entrena mediante PSO de manera de obtener los pesos de la red, para ello se utilizará un rango de 8 años de nuestra muestra de 10 años (un 80 de la muestra).
6. **Testeo de la red:** se obtienen los resultados una vez ya entrenada la red, para ello se utilizará un determinado número de datos, que serán el resto de los años no usados por el entrenamiento que corresponde a los últimos 2 años de la muestra.
7. **Comparación de resultados:** en la etapa final se compararán los resultados obtenidos por los diversos modelos mediante métricas especificadas más adelante.

6.3. Métricas de Calidad

6.3.1. Raíz Error Cuadrático Medio (RMSE)

Está dado por:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (d_i - y_i)^2}{N}} \quad (12)$$

Consiste en la suma de las diferencias entre los datos observados y los datos proyectados por el modelo.

6.3.2. Coeficiente de Determinación (R^2)

Está dado por:

$$R^2 = 1 - \frac{\sum_{i=1}^N (d_i - y_i)^2}{\sum_{i=1}^N (d_i - \bar{d})^2} \quad (13)$$

Mide la dependencia entre los datos reales y los pronosticados.

6.4. Entorno de Trabajo

El entorno de trabajo considera las características de hardware y software utilizadas para desarrollar e implementar los modelos.

Hardware:

- Procesador Intel I5, de 3.20GHZ
- Memoria RAM de 8 GB

Software:

- Windows 7 ultimate de 64 bits
- Matlab 2012b

7. Resultados Computacionales

A continuación se presentan los resultados obtenidos de las diversas configuraciones de la red PSO donde se comparará los resultados con LPSO. Cabe destacar que cada ejecución inicial fue con un número de 2000 iteraciones y 15 partículas de configuración de PSO que a medida que se obtienen las mejores configuraciones se testea con otros valores. Los valores de configuración de PSO son:

$$W_{max} = 0,9 \quad W_{min} = 0,2 \quad C_1 = 1,05 \quad C_2 = 2,95 \quad (14)$$

7.1. Análisis de los resultados con Modelo Tradicional PSO - Fallidos del Gran Valparaíso

A continuación se puede apreciar en la Tabla 2 los mejores resultados de todas las configuraciones probadas para todos los nodos de entrada y ocultos. Cabe mencionar que los datos son del resultado del Testing y se tomó valor de referencia el MSE obtenido de la etapa de Testing de los datos para seleccionar el resto de los indicadores.

Nodo Entrada	Nodo Oculto	MSE	RMSE	R	R2
1	5	0.019767172	0.140600692	0.497730453	0.247735604
2	15	0.010793971	0.106141998	0.757291753	0.5734908
3	9	0.00984434	0.102220353	0.776805066	0.603426111
4	15	0.009279169	0.103172356	0.772096467	0.596132954
5	4	0.006385127	0.07990699	0.871824554	0.760078054
6	2	0.006173311	0.07857042	0.87628272	0.767871405
7	8	0.006329208	0.081778311	0.865432454	0.748973332

8	2	0.004485608	0.066974685	0.911449845	0.83074082
9	10	0.004488053	0.067145554	0.910747536	0.829461074
10	11	0.004090644	0.073999774	0.892177792	0.795981212
11	6	0.003641024	0.061474851	0.925818102	0.857139158
12	2	0.003497579	0.062313377	0.918561475	0.843755184
13	9	0.003826878	0.061861764	0.919817082	0.846063465
14	9	0.003443779	0.058683719	0.92906895	0.863169114
15	5	0.004151301	0.06443059	0.914257461	0.835866704
16	7	0.003226636	0.056803485	0.933202276	0.870866488
17	2	0.003568154	0.059734026	0.927146651	0.859600912
18	5	0.004025913	0.063450083	0.917451038	0.841716407
19	2	0.005233658	0.072344022	0.895234303	0.801444458
20	5	0.004680107	0.068411308	0.909570786	0.827319015
21	2	0.004183208	0.064677727	0.912487361	0.832633185
22	10	0.00647722	0.080481176	0.862872303	0.744548611
23	2	0.005087016	0.071323318	0.896829288	0.804302772
24	2	0.006315998	0.079473253	0.870861732	0.758400157
25	10	0.007060019	0.084023918	0.850021916	0.722537258
26	3	0.006168485	0.078539704	0.872280866	0.760873909
27	2	0.004846853	0.069619347	0.899693291	0.809448018
28	4	0.00670601	0.081890229	0.86034985	0.740201864
29	6	0.007419645	0.08613736	0.882496422	0.778799934
30	1	0.009035926	0.09505749	0.882232882	0.778334857
31	3	0.004669446	0.068333347	0.905818742	0.820507594
32	1	0.010338282	0.101677344	0.84670255	0.716905209
33	4	0.0078336	0.088507626	0.829410697	0.687922104
34	2	0.007750699	0.088038055	0.833598308	0.694886139
35	5	0.009374671	0.096822883	0.78958859	0.623450141
36	2	0.009632837	0.098147018	0.790340758	0.624638514
37	2	0.009082308	0.095301146	0.799927578	0.639884131
38	6	0.008149131	0.090272539	0.828007819	0.685596948
39	2	0.008908669	0.094385744	0.825856949	0.6820397
40	4	0.009096717	0.095376712	0.797382591	0.635818996
41	2	0.010269077	0.101336456	0.794149186	0.63067293
42	8	0.011810312	0.108675261	0.727371262	0.529068952
43	4	0.009350701	0.096699022	0.818311229	0.669633267
44	12	0.011607576	0.107738459	0.735909906	0.54156339

45	3	0.010427437	0.102114824	0.777724608	0.604855566
46	4	0.011609084	0.107745461	0.75320601	0.567319294
47	9	0.01099631	0.104863292	0.752367211	0.56605642
48	2	0.00863106	0.092903501	0.817128741	0.667699379

Metricas					
Minimo	1	0.003226636	0.056803485	0.497730453	0.247735604
Maximo	15	0.019767172	0.140600692	0.933202276	0.870866488
Desviacion Estandar	3.6667	0.0032051	0.0179019	0.0783670	0.1215373
Promedio	5.2083	0.0073702	0.0845871	0.8444770	0.7191548

Tabla 2: PSO - Muertos Gran Valparaíso

Como se puede observar en los resultados de MSE la mejor configuración que se obtuvo para este set de datos fue de 16 nodos de entradas con 7 nodos ocultos. Lo cual nos indica que el modelo puede predecir con buenos resultados un desfase de 16 semanas. Los resultados nos demuestran que para este set de datos el modelo pudo predecir con un R^2 de 87%. Como se puede observar en el cuadro de métricas el promedio de los 48 nodos ocultos nos dio un R^2 de 71% el cual nos indica que para este set de datos el modelo predijo con buenos resultados para casi todos los nodos de entrada. Como se puede observar en la desviación estándar los valores son bastante bajos lo que nos demuestra que la variabilidad de los resultados no es tan amplia.

En la figura 10 se grafica los resultados obtenidos para la configuración que se obtuvo mejores métricas.

Como se puede apreciar en el gráfico de error relativo se puede observar,

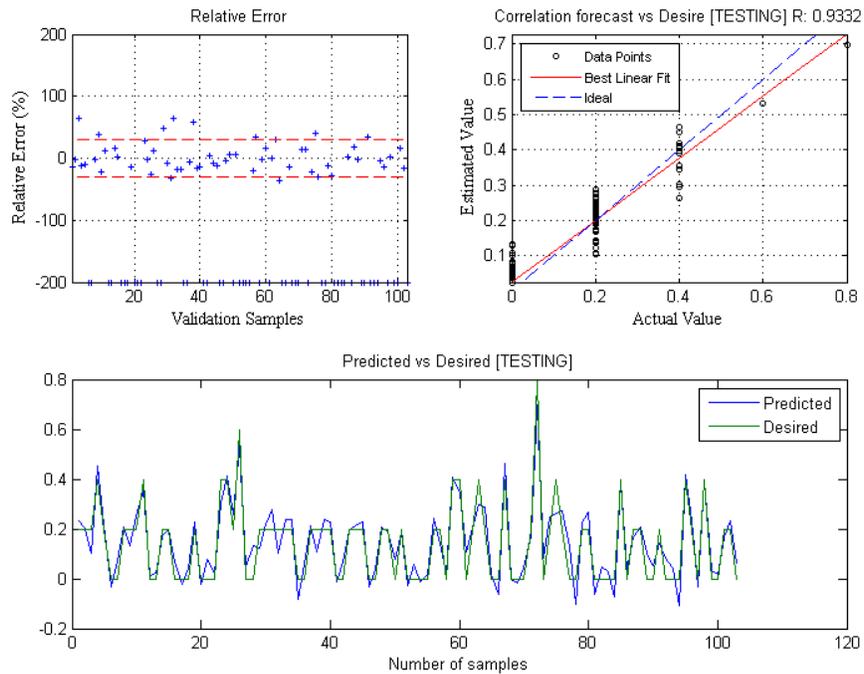


Figura 10: Grafico 16 nodos de entrada, 7 ocultos

que una importante fracción de la información analizada es aceptable debido a que los valores residuales son bajos, con una banda de ajuste que van desde -25% a 25% . En el gráfico de la correlación pronosticada podemos observar que la mayoría de los puntos se están agrupando en la cercanía de la recta ideal. Para el modelo la recta pronosticada se asemeja bastante a la recta ideal. Como se puede observar en el gráfico de la data pronosticada v/s deseada la gráfica representa una curva que se asemeja bastante a la data deseada a pesar de tener un R^2 que rodea el 86% .

7.2. Análisis de los resultados con Modelo Tradicional PSO - Lesionados del Gran Valparaíso

A continuación se puede apreciar en la Tabla 3 los mejores resultados de todas las configuraciones probadas para todos los nodos de entrada y ocultos. Cabe mencionar que los datos son del resultado del Testing y se tomó valor de referencia el MSE obtenido de la etapa de Testing de los datos para seleccionar el resto de los indicadores.

Nodo Entrada	Nodo Oculto	MSE	RMSE	R	R2
1	10	0.007153601	0.08457896	0.602265612	0.362723867
2	15	0.003401799	0.058324941	0.833775776	0.695182045
3	13	0.003375375	0.058097979	0.836059039	0.698994716
4	8	0.002848023	0.053366873	0.864260874	0.746946859
5	12	0.001951945	0.044180822	0.908663802	0.825669904
6	2	0.001997057	0.044688442	0.906588647	0.821902975
7	12	0.001731646	0.041613055	0.920654468	0.847604649
8	4	0.001474206	0.03839539	0.932438497	0.869441551
9	3	0.001386268	0.037232625	0.937315332	0.878560032
10	3	0.001318804	0.036315339	0.940193488	0.883963794
11	3	0.001313679	0.036244704	0.939690668	0.883018552
12	3	0.001236021	0.035157091	0.944149783	0.891418813
13	2	0.001211096	0.034800798	0.947396892	0.897560872
14	2	0.001663498	0.040786004	0.923740433	0.853296388
15	2	0.001501267	0.038746193	0.93177273	0.86820042
16	10	0.001746966	0.041796726	0.92313759	0.85218301
17	8	0.001729788	0.041590716	0.926052054	0.857572407
18	8	0.002263125	0.047572317	0.897003987	0.804616153
19	3	0.001844744	0.042950479	0.917141803	0.841149087
20	2	0.00206494	0.045441611	0.907924127	0.82432622
21	2	0.002505868	0.050058646	0.884972713	0.783176703
22	5	0.002095531	0.045776969	0.904194053	0.817566886
23	9	0.002620088	0.051186792	0.881312477	0.776711682

24	5	0.002223614	0.047155211	0.898581358	0.807448456
25	7	0.002880173	0.05366724	0.866293942	0.750465194
26	4	0.002904182	0.053890459	0.867818205	0.753108436
27	1	0.002347464	0.04845063	0.912320209	0.832328163
28	5	0.00293509	0.054176467	0.8640031	0.746501357
29	1	0.002647947	0.051458207	0.894071028	0.799363003
30	15	0.00326098	0.057104989	0.857593216	0.735466125
31	3	0.002951812	0.054330578	0.869187491	0.755486894
32	9	0.003244881	0.056963857	0.85958155	0.738880441
33	2	0.003445165	0.058695531	0.840617832	0.706638339
34	1	0.003621417	0.060178214	0.862353866	0.74365419
35	2	0.002984611	0.054631592	0.863417687	0.745490102
36	1	0.003468741	0.058896016	0.840901556	0.707115427
37	1	0.004014152	0.063357335	0.822059916	0.675782506
38	3	0.003922702	0.062631475	0.825078181	0.680754004
39	2	0.004043377	0.063587553	0.810280675	0.656554772
40	4	0.00409821	0.064017262	0.810565137	0.657015841
41	1	0.004427924	0.066542645	0.830455623	0.689656542
42	1	0.004929907	0.070213294	0.770769591	0.594085763
43	1	0.00370781	0.060891793	0.840958472	0.707211152
44	13	0.003410597	0.058400318	0.848317441	0.71964248
45	5	0.004291413	0.065508881	0.820960688	0.673976451
46	10	0.003606911	0.060057567	0.832963887	0.693828836
47	1	0.00417471	0.064611995	0.811183454	0.658018596
48	1	0.005089827	0.071343023	0.752543008	0.566320979

Metricas					
Minimo	1	0.001211096	0.034800798	0.602265612	0.362723867
Maximo	15	0.007153601	0.08457896	0.947396892	0.897560872
Desviacion Estandar	4,2174	0,00121497	0,01106279	0,06157980	0,10077612
Promedio	5,0	0,00289727	0,05270137	0,86840796	0,75784545

Tabla 3: PSO - Lesionados Gran Valparaíso

Como se puede observar en los resultados de MSE la mejor configuración

que se obtuvo para este set de datos fue de 12 nodos de entradas con 3 nodos ocultos. Lo cual nos indica que el modelo puede predecir con buenos resultados un desfase de 12 semanas. Los resultados nos demuestran que para este set de datos el modelo pudo predecir con un R^2 de 89%. Como se puede observar en el cuadro de métricas el promedio de los 48 nodos ocultos nos dio un R^2 de 75% el cual nos indica que para este set de datos el modelo predijo con buenos resultados para casi todos los nodos de entrada. Como se puede observar en la desviación estándar los valores son bastante bajos lo que nos demuestra que la variabilidad de los resultados no es tan amplia.

En la figura 11 se grafica los resultados obtenidos para la configuración de 12 nodos de entradas y 3 nodos de salida, donde se obtuvieron mejores métricas.

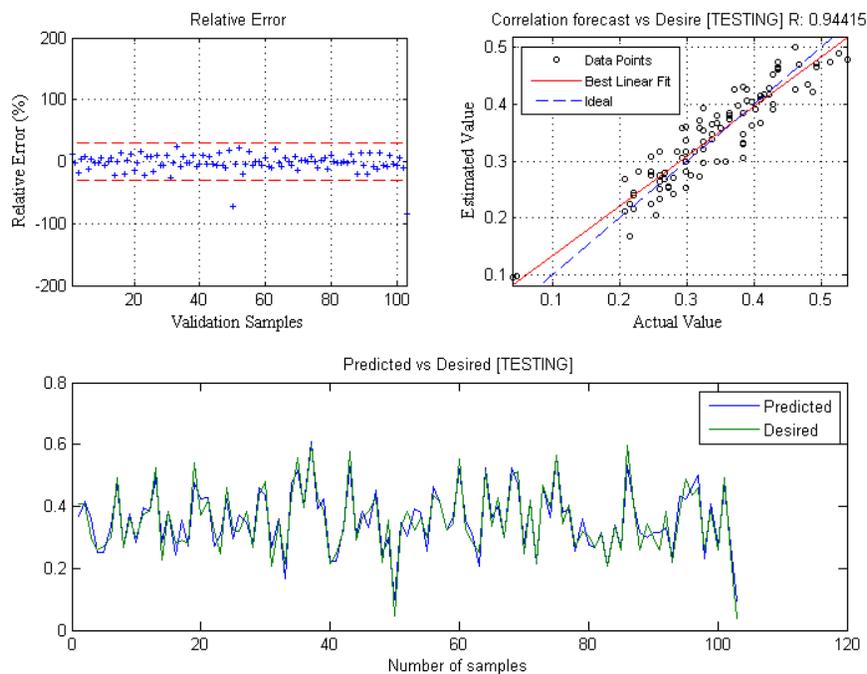


Figura 11: Grafico 12 nodos de entrada, 3 ocultos

Como se puede apreciar en el gráfico de error relativo se puede observar, que una importante fracción de la información analizada es aceptable debido a que los valores residuales son bajos, con una banda de ajuste que van desde -25% a 25% . En el gráfico de la correlación pronosticada podemos observar que la mayoría de los puntos se están agrupando en la cercanía de la recta ideal. Para el modelo la recta pronosticada se asemeja bastante a la recta ideal. Como se puede observar en el gráfico de la data pronosticada v/s deseada la gráfica representa una curva que se asemeja bastante a la data deseada a pesar de tener un R^2 que rodea el 89% .

7.3. Análisis de los resultados con Modelo Tradicional PSO - Muertos Otras Comunas

A continuación se puede apreciar en la Tabla 4 los mejores resultados de todas las configuraciones probadas para todos los nodos de entrada y ocultos. Cabe mencionar que los datos son del resultado del Testing y se tomó valor de referencia el MSE obtenido de la etapa de Testing de los datos para seleccionar el resto de los indicadores.

Nodo Entrada	Nodo Oculto	MSE	RMSE	R	R2
1	2	0.022309138	0.149362437	0.560236421	0.313864848
2	3	0.008026326	0.089589766	0.871228913	0.759039819
3	5	0.007980454	0.089333387	0.869269516	0.755629492
4	4	0.008246655	0.090811096	0.864355923	0.747111162
5	4	0.006633661	0.081447292	0.892422046	0.796417108
6	13	0.00590346	0.07683398	0.905256409	0.819489167
7	13	0.006289616	0.079307099	0.899103874	0.808387777
8	7	0.005117184	0.071534498	0.919247573	0.845016101

9	15	0.00494064	0.070289689	0.922891886	0.851729433
10	14	0.004982695	0.070588205	0.921406458	0.848989861
11	2	0.004559882	0.067526899	0.928026962	0.861234043
12	7	0.004204625	0.064843077	0.929795402	0.86451949
13	3	0.003985874	0.063133781	0.933417347	0.871267943
14	2	0.004327827	0.065786225	0.927370675	0.860016369
15	8	0.005248679	0.072447769	0.910278981	0.828607823
16	6	0.005839976	0.076419735	0.901272138	0.812291467
17	4	0.004979333	0.070564389	0.916157699	0.839344929
18	2	0.005602428	0.074849366	0.90690934	0.822484551
19	3	0.006124633	0.078260032	0.89618237	0.80314284
20	10	0.006888091	0.082994524	0.889476799	0.791168975
21	4	0.005481042	0.074034063	0.907833689	0.824162006
22	2	0.005594843	0.074798684	0.904135369	0.817460765
23	3	0.008210626	0.090612502	0.862277053	0.743521716
24	3	0.007309057	0.085493026	0.874540052	0.764820302
25	6	0.007126942	0.084421215	0.876128296	0.767600791
26	5	0.00921292	0.095983957	0.836798115	0.700231085
27	3	0.007178606	0.084726656	0.875054318	0.765720059
28	2	0.008383713	0.091562619	0.85327481	0.728077901
29	10	0.009149764	0.095654399	0.849434775	0.721539437
30	1	0.010168778	0.100840359	0.838263002	0.702684861
31	1	0.009841408	0.099203869	0.843660224	0.711762573
32	8	0.010575059	0.102835104	0.810740702	0.657300486
33	8	0.008794393	0.093778424	0.857463137	0.735243031
34	6	0.011832776	0.108778564	0.786259086	0.61820335
35	10	0.01072531	0.10356307	0.81524354	0.66462203
36	1	0.010507073	0.102504014	0.82690301	0.683768588
37	3	0.009385169	0.096877083	0.835952714	0.698816941
38	2	0.009804202	0.099016169	0.825748494	0.681860575
39	6	0.010287886	0.101429219	0.818031075	0.669174839
40	1	0.010542618	0.10267725	0.823835216	0.678704463
41	2	0.012891584	0.11354111	0.767722407	0.589397694
42	14	0.013046584	0.114221644	0.757332766	0.573552918
43	1	0.011250016	0.106066091	0.816487259	0.666651444
44	1	0.011038081	0.105062271	0.81353426	0.661837992
45	1	0.011026407	0.105006701	0.811003975	0.657727447

46	1	0.009549469	0.097721384	0.84130086	0.707787138
47	6	0.012545969	0.112008791	0.769923883	0.592782785
48	2	0.012855451	0.11338188	0.757224881	0.573389521

Metricas					
Minimo	1	0.003985874	0.063133781	0.560236421	0.313864848
Maximo	15	0.022309138	0.149362437	0.933417347	0.871267943
Desviacion Estandar	3,9840	0,00332806	0,01712711	0,06562961	0,10431901
Promedio	5,0	0,00846889	0,09045257	0,85459195	0,73454492

Tabla 4: PSO - Muertos Otras Comunas

Como se puede observar en los resultados de MSE la mejor configuración que se obtuvo para este set de datos fue de 13 nodos de entradas con 3 nodos ocultos. Lo cual nos indica que el modelo puede predecir con buenos resultados un desfase de 13 semanas. Los resultados nos demuestran que para este set de datos el modelo pudo predecir con un R^2 de 87%. Como se puede observar en el cuadro de métricas el promedio de los 48 nodos ocultos nos dio un R^2 de 73% el cual nos indica que para este set de datos el modelo predijo con buenos resultados para casi todos los nodos de entrada. Como se puede observar en la desviación estándar los valores son bastantes bajos lo que nos demuestra que la variabilidad de los resultados no es tan amplia.

En la figura 12 se grafica los resultados obtenidos para la configuración de 13 nodos de entradas y 3 nodos de salida, donde se obtuvieron mejores métricas.

Como se puede apreciar en el gráfico de error relativo se puede observar, que una importante fracción de la información analizada es aceptable debido a que los valores residuales son bajos, con una banda de ajuste que van desde

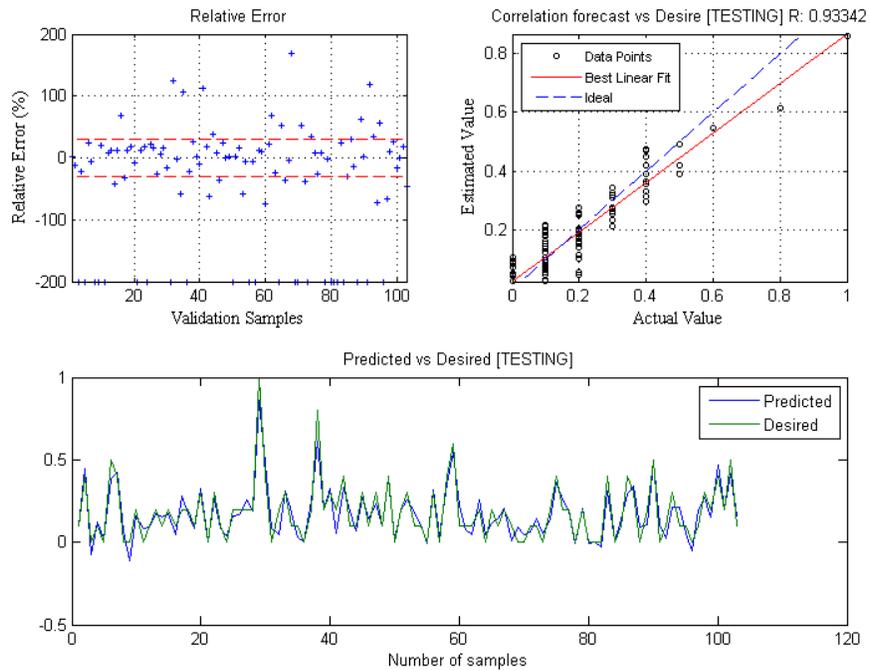


Figura 12: Grafico 13 nodos de entrada, 3 ocultos

−25 % a 25 %. En el gráfico de la correlación pronosticada podemos observar que la mayoría de los puntos se están agrupando en la cercanía de la recta ideal. Para el modelo la recta pronosticada se asemeja bastante a la recta ideal. Como se puede observar en el gráfico de la data pronosticada v/s deseada la gráfica representa una curva que se asemeja bastante a la data deseada a pesar de tener un R^2 que rodea el 87 %.

7.4. Análisis de los resultados con Modelo Tradicional PSO - Lesionados Otras Comunas

A continuación se puede apreciar en la Tabla 5 los mejores resultados de todas las configuraciones probadas para todos los nodos de entrada y ocultos. Cabe mencionar que los datos son del resultado del Testing y se tomó valor de

referencia el MSE obtenido de la etapa de Testing de los datos para seleccionar el resto de los indicadores.

Nodo Entrada	Nodo Oculto	MSE	RMSE	R	R2
1	11	0.010353566	0.101752473	0.654993915	0.429017029
2	15	0.006047583	0.077766208	0.821848563	0.675435061
3	2	0.005360351	0.073214417	0.838817268	0.703614408
4	14	0.00508513	0.071310096	0.844330454	0.712893915
5	14	0.003973471	0.063035478	0.88218787	0.778255438
6	14	0.003890038	0.062370169	0.883983173	0.78142625
7	8	0.003683249	0.060689776	0.89271925	0.796947659
8	12	0.003170484	0.056307052	0.912795289	0.833195239
9	4	0.003134584	0.055987356	0.90986495	0.827854228
10	2	0.003009979	0.054863278	0.912837839	0.83327292
11	10	0.003157619	0.056192697	0.911983872	0.831714582
12	9	0.00307907	0.055489372	0.923641417	0.853113467
13	2	0.003495414	0.059122027	0.902628928	0.814738982
14	4	0.002944691	0.054265006	0.913150678	0.833844161
15	4	0.003417617	0.058460391	0.907494684	0.823546601
16	3	0.003284282	0.057308658	0.905257464	0.819491076
17	2	0.004806991	0.06933247	0.877194153	0.769469581
18	2	0.00246663	0.049665176	0.936522214	0.877073858
19	4	0.00342928	0.058560054	0.901560403	0.812811161
20	15	0.004786507	0.069184586	0.872405999	0.761092227
21	9	0.004988809	0.070631499	0.857280534	0.734929914
22	2	0.00400782	0.063307342	0.88694094	0.786664232
23	4	0.005551618	0.074509178	0.831994896	0.692215507
24	11	0.005758223	0.075882953	0.824288996	0.679452348
25	3	0.005981292	0.077338812	0.817900193	0.668960726
26	2	0.00501676	0.070829091	0.847069899	0.717527414
27	3	0.005790869	0.076097756	0.822182877	0.675984684
28	7	0.005754718	0.075859856	0.817859674	0.668894446
29	6	0.007136577	0.084478262	0.771868133	0.595780415
30	8	0.006151123	0.078429097	0.811748919	0.658936308
31	2	0.006064469	0.0778747	0.806882797	0.651059848
32	5	0.006935733	0.083281047	0.770747339	0.594051461

33	5	0.007053297	0.083983907	0.768894615	0.591198929
34	2	0.006924574	0.083214026	0.796282668	0.634066087
35	10	0.007671526	0.087587246	0.743778798	0.553206901
36	4	0.006837651	0.082690088	0.777686922	0.604796949
37	2	0.007544094	0.086856747	0.761009948	0.57913614
38	2	0.006634207	0.081450641	0.794013285	0.630457097
39	3	0.006872362	0.08289971	0.796299561	0.634092992
40	4	0.007790943	0.088266319	0.749974299	0.562461448
41	7	0.008375478	0.091517639	0.721448802	0.520488374
42	12	0.008581213	0.092634837	0.726633422	0.52799613
43	2	0.008136711	0.090203719	0.742345109	0.551076261
44	15	0.009693254	0.098454325	0.682755486	0.466155053
45	2	0.009390176	0.096902919	0.704698332	0.496599739
46	10	0.007305095	0.085469846	0.768943663	0.591274357
47	2	0.008937938	0.094540669	0.693463544	0.480891687
48	2	0.010561738	0.102770319	0.657124565	0.431812694

Metricas					
Minimo	2	0.00246663	0.049665176	0.654993915	0.429017029
Maximo	15	0.010561738	0.102770319	0.936522214	0.877073858
Desviacion Estandar	4,4657	0,00216398	0,01429085	0,07670112	0,12366327
Promedio	6,1875	0,00583385	0,07505915	0,81996535	0,67810367

Tabla 5: PSO - Lesionados Otras Comunas

Como se puede observar en los resultados de MSE la mejor configuración que se obtuvo para este set de datos fue de 18 nodos de entradas con 2 nodos ocultos. Lo cual nos indica que el modelo puede predecir con buenos resultados un desfase de 18 semanas. Los resultados nos demuestran que para este set de datos el modelo pudo predecir con un R^2 de 87%. Como se puede observar en el cuadro de métricas el promedio de los 48 nodos ocultos nos dio un R^2 de 67% el cual nos indica que para este set de datos el modelo predijo con buenos

resultados para casi todos los nodos de entrada. Como se puede observar en la desviación estándar los valores son bastantes bajos lo que nos demuestra que la variabilidad de los resultados no es tan amplia.

En la figura 13 se grafica los resultados obtenidos para la configuración de 18 nodos de entradas y 2 nodos de salida, donde se obtuvieron mejores métricas.

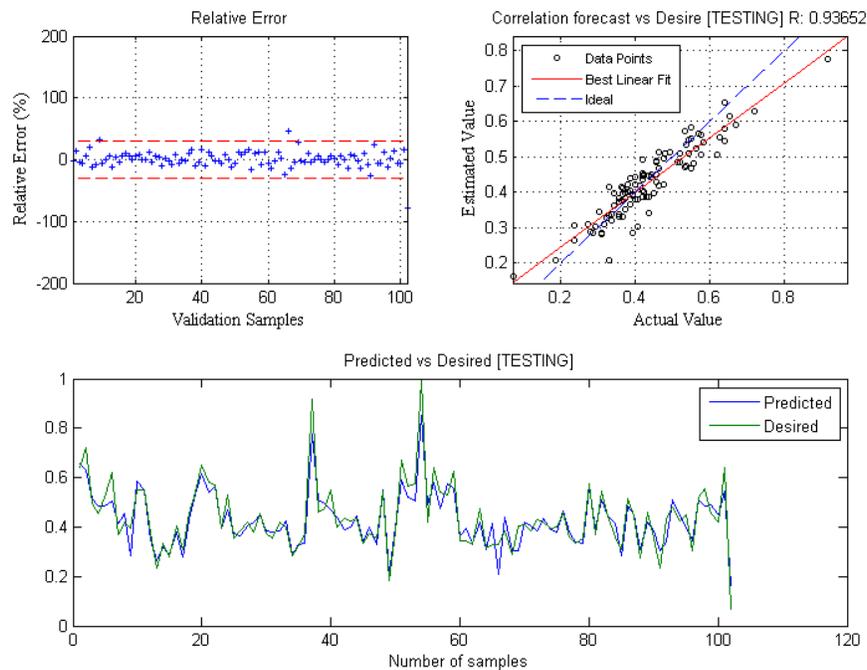


Figura 13: Grafico 18 nodos de entrada, 2 ocultos

Como se puede apreciar en el gráfico de error relativo se puede observar, que una importante fracción de la información analizada es aceptable debido a que los valores residuales son bajos, con una banda de ajuste que van desde -25% a 25% . En el gráfico de la correlación pronosticada podemos observar que la mayoría de los puntos se están agrupando en la cercanía de la recta ideal. Para el modelo la recta pronosticada se asemeja bastante a la recta ideal. Como

se puede observar en el gráfico de la data pronosticada v/s deseada la gráfica representa una curva que se asemeja bastante a la data deseada a pesar de tener un R^2 que rodea el 87 %.

7.5. Análisis de los resultados con Modelo PSO con Búsqueda Local Adaptativa - Fallecidos del Gran Valparaíso

A continuación se puede apreciar en la Tabla 6 los mejores resultados de todas las configuraciones probadas para todos los nodos de entrada y ocultos. Cabe mencionar que los datos son del resultado del Testing y se tomó valor de referencia el MSE obtenido de la etapa de Testing de los datos para seleccionar el resto de los indicadores.

Nodo Entrada	Nodo Oculto	MSE	RMSE	R	R2
1	7	0.019767781	0.140597942	0.497741982	0.24774708
2	13	0.010640226	0.10315147	0.771936704	0.595886276
3	6	0.009808339	0.099037058	0.793295853	0.629318311
4	4	0.009448298	0.097202354	0.805474141	0.648788592
5	4	0.006704807	0.081882883	0.865163462	0.748507816
6	5	0.006372391	0.079827257	0.873385879	0.762802894
7	2	0.006144719	0.078388258	0.875760357	0.766956203
8	9	0.004313332	0.065675959	0.915312812	0.837797543
9	10	0.004643388	0.068142408	0.909317705	0.826858689
10	15	0.004405724	0.066375631	0.913477436	0.834441026
11	3	0.003434093	0.058601131	0.933240297	0.870937451
12	2	0.00352056	0.059334304	0.926253525	0.857945593
13	2	0.003494081	0.059110754	0.927277994	0.859844478
14	2	0.003842274	0.061986075	0.921964557	0.850018645
15	4	0.004271562	0.065357187	0.909896383	0.827911427
16	10	0.003905585	0.062494678	0.91932816	0.845164265
17	4	0.004211295	0.064894489	0.91389238	0.835199282
18	2	0.003694449	0.06078198	0.926728222	0.858825198

19	2	0.005502634	0.07417974	0.886485643	0.785856794
20	4	0.003454327	0.058773526	0.929784483	0.864499184
21	3	0.005036183	0.070966068	0.903484969	0.816285089
22	3	0.005308604	0.072860169	0.888230658	0.788953702
23	2	0.005191315	0.07205078	0.893049266	0.797536992
24	3	0.006465566	0.080408743	0.867573594	0.752683941
25	2	0.007013887	0.083748954	0.84968858	0.721970683
26	2	0.006074412	0.077938511	0.872913853	0.761978594
27	2	0.004936896	0.070263046	0.897115004	0.804815331
28	2	0.005830713	0.076359104	0.881175513	0.776470284
29	3	0.006679542	0.081728463	0.862166604	0.743331252
30	6	0.006611375	0.081310364	0.860105173	0.739780908
31	5	0.009192069	0.095875279	0.798578565	0.637727724
32	12	0.007576669	0.087044066	0.840851721	0.707031617
33	2	0.01037407	0.101853177	0.784603693	0.615602954
34	6	0.009481223	0.097371571	0.789105764	0.622687907
35	5	0.009483297	0.097382221	0.792557532	0.628147442
36	8	0.010766358	0.103761064	0.844051347	0.712422676
37	10	0.009822416	0.099108103	0.78351031	0.613888405
38	3	0.009691645	0.098446152	0.785085263	0.61635887
39	4	0.008724343	0.093404194	0.815943431	0.665763682
40	8	0.011132471	0.105510526	0.759841875	0.577359675
41	4	0.010924091	0.104518376	0.759236836	0.576440573
42	2	0.009828486	0.099138723	0.789419953	0.623183862
43	4	0.010591429	0.10291467	0.770735227	0.594032791
44	4	0.010779727	0.103825466	0.761112424	0.579292123
45	3	0.009181894	0.0958222	0.797573215	0.636123033
46	5	0.010077127	0.100384892	0.777376004	0.604313452
47	7	0.010385997	0.101911714	0.769423844	0.592013052
48	9	0.014267305	0.119445824	0.675593231	0.456426214

Metricas					
Minimo	2	0.003434093	0.058601131	0.497741982	0.24774708
Maximo	15	0.019767781	0.140597942	0.933240297	0.870937451
Desviacion Estandar	3.2713	0.0033188	0.0184586	0.0807577	0.1255993

Promedio	4.9791	0.0075627	0.0850239	0.8392879	0.7107902
----------	--------	-----------	-----------	-----------	-----------

Tabla 6: LPSO - Muertos Gran Valparaíso

Como se puede observar en los resultados de MSE la mejor configuración que se obtuvo para este set de datos fue de 11 nodos de entradas con 3 nodos ocultos. Lo cual nos indica que el modelo puede predecir con buenos resultados un desfase de 11 semanas. Los resultados nos demuestran que para este set de datos el modelo pudo predecir con un R^2 de 87 %. Como se puede observar en el cuadro de métricas el promedio de los 48 nodos ocultos nos dio un R^2 de 71 % el cual nos indica que para este set de datos el modelo predijo con buenos resultados para casi todos los nodos de entrada. Como se puede observar en la desviación estándar los valores son bastante bajos lo que nos demuestra que la variabilidad de los resultados no es tan amplia.

En la figura 14 se grafica los resultados obtenidos para la configuración de 11 nodos de entradas y 3 nodos de salida, donde se obtuvieron mejores métricas.

Como se puede apreciar en el gráfico de error relativo se puede observar, que una importante fracción de la información analizada es aceptable debido a que los valores residuales son bajos, con una banda de ajuste que van desde -25% a 25% . En el gráfico de la correlación pronosticada podemos observar que la mayoría de los puntos se están agrupando en la cercanía de la recta ideal. Para el modelo la recta pronosticada se asemeja bastante a la recta ideal. Como se puede observar en el gráfico de la data pronosticada v/s deseada la gráfica representa una curva que se asemeja bastante a la data deseada a pesar de tener un R^2 que rodea el 87 %.

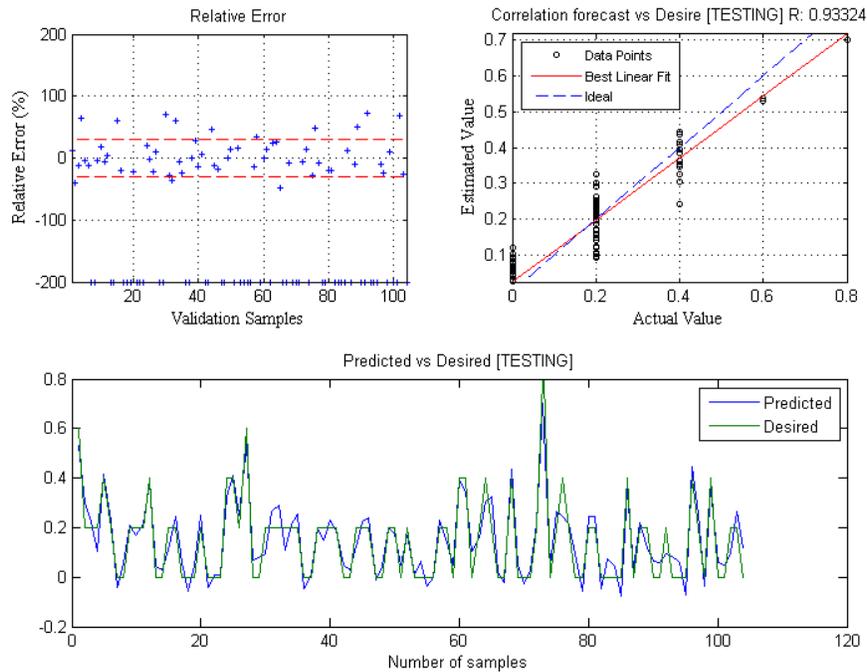


Figura 14: Grafico 11 nodos de entrada, 3 ocultos

7.6. Análisis de los resultados con Modelo PSO con Búsqueda Local Adaptativa - Lesionados del Gran Valparaíso

A continuación se puede apreciar en la Tabla 7 los mejores resultados de todas las configuraciones probadas para todos los nodos de entrada y ocultos. Cabe mencionar que los datos son del resultado del Testing y se tomó valor de referencia el MSE obtenido de la etapa de Testing de los datos para seleccionar el resto de los indicadores.

Nodo Entrada	Nodo Oculto	MSE	RMSE	R	R2
1	9	0.00715321	0.084576652	0.602226869	0.362677202
2	7	0.003381302	0.058148966	0.835042289	0.697295624
3	8	0.003415977	0.058446358	0.832838147	0.693619378

4	9	0.00281806	0.053085404	0.866033767	0.750014486
5	2	0.001909683	0.04369992	0.911470232	0.830777984
6	12	0.001980808	0.04450627	0.90713524	0.822894344
7	5	0.00172219	0.041499275	0.920484925	0.847292496
8	2	0.001465595	0.038283088	0.933545972	0.871508082
9	7	0.001492099	0.038627698	0.931097394	0.866942356
10	7	0.001423271	0.037726264	0.934562473	0.873407016
11	6	0.001322298	0.036363412	0.939303872	0.882291763
12	3	0.001597438	0.039967959	0.93376964	0.87192574
13	3	0.0012366	0.035165324	0.944121402	0.891365221
14	5	0.001431087	0.037829707	0.942739426	0.888757626
15	9	0.001750078	0.041833936	0.920696174	0.847681444
16	4	0.001647675	0.040591559	0.927753579	0.860726703
17	3	0.00182866	0.042762829	0.916421454	0.839828281
18	6	0.001540483	0.039248981	0.930018878	0.864935114
19	3	0.002881325	0.053677976	0.874939733	0.765519537
20	4	0.002134385	0.0461994	0.902439897	0.814397768
21	3	0.002207438	0.046983385	0.900125769	0.8102264
22	2	0.00224644	0.049774927	0.888250793	0.788989472
23	3	0.002525036	0.050249738	0.883348091	0.78030385
24	3	0.002006905	0.044798493	0.909355081	0.826926664
25	5	0.003200656	0.056574339	0.85973427	0.739143015
26	2	0.002648995	0.051468386	0.877942014	0.77078218
27	5	0.00261413	0.051128561	0.88498362	0.783196007
28	10	0.00365263	0.06043699	0.828628374	0.686624982
29	2	0.002720617	0.05215953	0.879444183	0.773422072
30	3	0.003314299	0.057569953	0.85036849	0.723126569
31	2	0.003427037	0.058540903	0.83946976	0.704709477
32	1	0.003183106	0.056419021	0.871221984	0.759027746
33	4	0.003315087	0.057576789	0.850952641	0.724120397
34	1	0.003127232	0.055921656	0.872403635	0.761088102
35	6	0.003650606	0.060420242	0.83510907	0.697407159
36	2	0.003431626	0.058580078	0.841588996	0.708272038
37	7	0.00432176	0.065740092	0.794888146	0.631847165
38	2	0.003819837	0.061804832	0.826528065	0.683148643
39	1	0.004060215	0.063719816	0.814604221	0.663580037
40	6	0.003385156	0.058182096	0.847957038	0.719031139

41	4	0.003668925	0.060571652	0.82812575	0.685792258
42	2	0.004899853	0.06999895	0.765064955	0.585324386
43	2	0.003879873	0.062288623	0.820819343	0.673744394
44	3	0.003945724	0.062814996	0.814595945	0.663566554
45	2	0.003620475	0.060170382	0.837743067	0.701813445
46	10	0.004692463	0.068501557	0.791271513	0.626110608
47	15	0.004165652	0.064541866	0.806783116	0.650898997
48	1	0.003626233	0.060218212	0.842456146	0.709732357

Metricas					
Minimo	1	0.0012366	0.035165324	0.602226869	0.362677202
Maximo	15	0.00715321	0.084576652	0.944121402	0.891365221
Desviacion Estandar	3.1722	0.0011746	0.0106958	0.0608031	0.0995981
Promedio	4.6458	0.0029060	0.0529041	0.8660501	0.7536628

Tabla 7: LPSO - Lesionados Gran Valparaíso

Como se puede observar en los resultados de MSE la mejor configuración que se obtuvo para este set de datos fue de 13 nodos de entradas con 3 nodos ocultos. Lo cual nos indica que el modelo puede predecir con buenos resultados un desfase de 13 semanas. Los resultados nos demuestran que para este set de datos el modelo pudo predecir con un R^2 de 89%. Como se puede observar en el cuadro de métricas el promedio de los 48 nodos ocultos nos dio un R^2 de 75% el cual nos indica que para este set de datos el modelo predijo con buenos resultados para casi todos los nodos de entrada. Como se puede observar en la desviación estándar los valores son bastantes bajos lo que nos demuestra que la variabilidad de los resultados no es tan amplia.

En la figura 15 se grafica los resultados obtenidos para la configuración de 13 nodos de entradas y 3 nodos de salida, donde se obtuvieron mejores métricas.

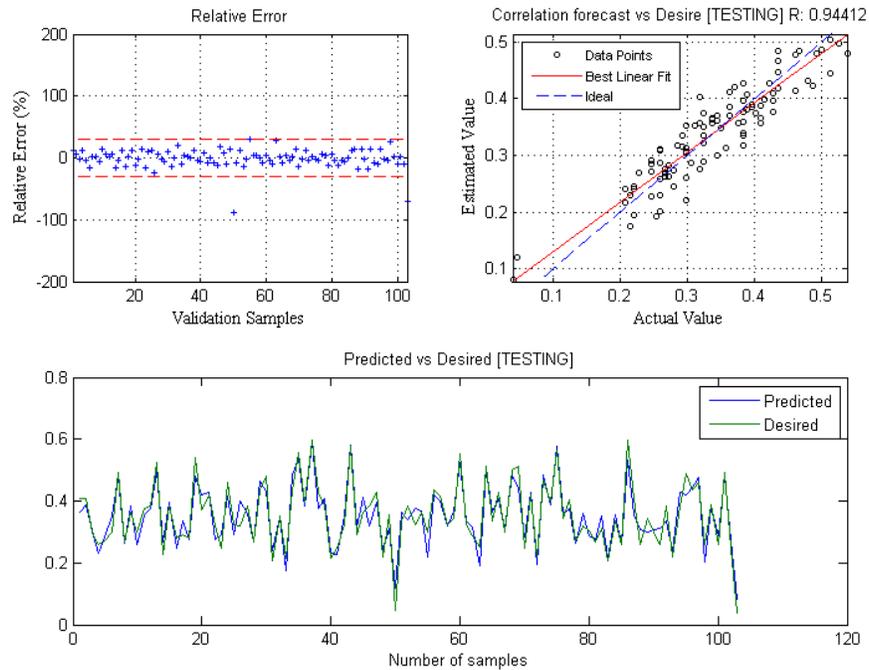


Figura 15: Grafico 13 nodos de entrada, 3 ocultos

Como se puede apreciar en el gráfico de error relativo se puede observar, que una importante fracción de la información analizada es aceptable debido a que los valores residuales son bajos, con una banda de ajuste que van desde -25% a 25% . En el gráfico de la correlación pronosticada podemos observar que la mayoría de los puntos se están agrupando en la cercanía de la recta ideal. Para el modelo la recta pronosticada se asemeja bastante a la recta ideal. Como se puede observar en el gráfico de la data pronosticada v/s deseada la gráfica representa una curva que se asemeja bastante a la data deseada a pesar de tener un R^2 que rodea el 89% .

7.7. Análisis de los resultados con Modelo PSO con Búsqueda Local Adaptativa - Fallecidos Otras Comunas

A continuación se puede apreciar en la Tabla 8 los mejores resultados de todas las configuraciones probadas para todos los nodos de entrada y ocultos. Cabe mencionar que los datos son del resultado del Testing y se tomó valor de referencia el MSE obtenido de la etapa de Testing de los datos para seleccionar el resto de los indicadores.

Nodo Entrada	Nodo Oculto	MSE	RMSE	R	R2
1	8	0.022295706	0.149317467	0.561005764	0.314727467
2	6	0.00811555	0.090086345	0.869530638	0.756083531
3	3	0.008436899	0.091852594	0.861393617	0.741998964
4	3	0.008094518	0.089969537	0.867532237	0.752612183
5	12	0.006715316	0.081947031	0.891044122	0.793959627
6	5	0.005714889	0.075596888	0.908251215	0.82492027
7	12	0.006037273	0.077699893	0.90353803	0.816380972
8	9	0.004961958	0.070441164	0.92318668	0.852273645
9	5	0.004730349	0.068777534	0.926647618	0.858675808
10	7	0.005023073	0.070873638	0.92050778	0.847334574
11	12	0.004389907	0.066256372	0.932330638	0.869240419
12	3	0.004438378	0.066621153	0.925350663	0.856273849
13	12	0.004927668	0.070197353	0.916529628	0.840026558
14	2	0.00459423	0.067780752	0.921820348	0.849752754
15	6	0.005007973	0.070767034	0.916696578	0.840332617
16	3	0.00497387	0.070525668	0.917842722	0.842435262
17	11	0.006565369	0.081026963	0.887041831	0.786843209
18	5	0.004672379	0.0683548	0.921279365	0.848755668
19	4	0.00576168	0.075905731	0.905080656	0.819170994
20	2	0.007797181	0.088301645	0.864439965	0.747256453
21	3	0.005967759	0.077251271	0.898014154	0.806429421
22	4	0.006130913	0.07830015	0.896210754	0.803193716
23	2	0.007081018	0.084148783	0.881063782	0.776273387

24	4	0.009299385	0.09643332	0.837300643	0.701072367
25	5	0.007780228	0.088205602	0.863930677	0.746376215
26	14	0.008448067	0.091913366	0.852063389	0.726012018
27	4	0.007839663	0.088541869	0.863429544	0.745510578
28	2	0.009167221	0.095745607	0.840137564	0.705831126
29	4	0.008538033	0.092401479	0.852531945	0.726810717
30	4	0.010484633	0.102394499	0.811863419	0.659122212
31	1	0.010914203	0.104471061	0.813372988	0.661575618
32	11	0.01005576	0.100278411	0.821485944	0.674839156
33	3	0.007864831	0.088683885	0.863859723	0.746253621
34	1	0.011600781	0.107706923	0.807572581	0.652173474
35	4	0.010936863	0.104579458	0.805684777	0.64912796
36	2	0.009951026	0.09975483	0.824853297	0.680382962
37	1	0.009841251	0.099203082	0.847025267	0.717451802
38	1	0.010544977	0.10268874	0.828444866	0.686320896
39	1	0.011704056	0.108185286	0.803091707	0.644956291
40	1	0.012562128	0.112080899	0.778012304	0.605303145
41	1	0.011105772	0.105383926	0.819303865	0.671258823
42	2	0.011275019	0.10618389	0.792870991	0.628644408
43	2	0.012558521	0.112064807	0.770067971	0.593004679
44	1	0.011796894	0.108613506	0.788494006	0.621722798
45	1	0.012748385	0.112908749	0.783525568	0.613912316
46	1	0.011404689	0.106792737	0.798479886	0.637570128
47	1	0.010811906	0.103980315	0.815474916	0.664999339
48	3	0.013142414	0.114640369	0.75651805	0.572319561

Metricas					
Minimo	1	0.004389907	0.066256372	0.561005764	0.314727467
Maximo	14	0.022295706	0.149317467	0.932330638	0.869240419
Desviacion Estandar	3.6898	0.00336505	0.01730190	0.06567002	0.10470413
Promedio	4.4583	0.00864189	0.09137159	0.85116114	0.72869799

Tabla 8: LPSO - Fallecidos Otras Comunas

Como se puede observar en los resultados de MSE la mejor configuración

que se obtuvo para este set de datos fue de 11 nodos de entradas con 12 nodos ocultos. Lo cual nos indica que el modelo puede predecir con buenos resultados un desfase de 11 semanas. Los resultados nos demuestran que para este set de datos el modelo pudo predecir con un R^2 de 86 %. Como se puede observar en el cuadro de métricas el promedio de los 48 nodos ocultos nos dio un R^2 de 72 % el cual nos indica que para este set de datos el modelo predijo con buenos resultados para casi todos los nodos de entrada. Como se puede observar en la desviación estándar los valores son bastantes bajos lo que nos demuestra que la variabilidad de los resultados no es tan amplia.

En la figura 16 se grafica los resultados obtenidos para la configuración de 11 nodos de entradas y 12 nodos de salida, donde se obtuvieron mejores métricas.

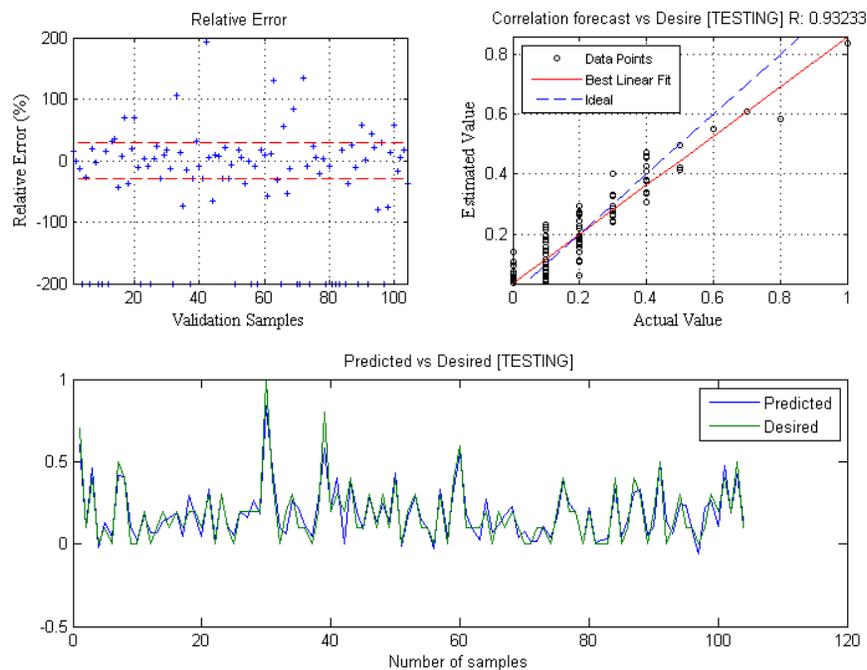


Figura 16: Grafico 11 nodos de entrada, 12 ocultos

Como se puede apreciar en el gráfico de error relativo se puede observar, que una importante fracción de la información analizada es aceptable debido a que los valores residuales son bajos, con una banda de ajuste que van desde -25% a 25% . En el gráfico de la correlación pronosticada podemos observar que la mayoría de los puntos se están agrupando en la cercanía de la recta ideal. Para el modelo la recta pronosticada se asemeja bastante a la recta ideal. Como se puede observar en el gráfico de la data pronosticada v/s deseada la gráfica representa una curva que se asemeja bastante a la data deseada a pesar de tener un R^2 que rodea el 86% .

7.8. Análisis de los resultados con Modelo PSO con Búsqueda Local Adaptativa - Lesionados Otras Comunas

A continuación se puede apreciar en la Tabla 9 los mejores resultados de todas las configuraciones probadas para todos los nodos de entrada y ocultos. Cabe mencionar que los datos son del resultado del Testing y se tomó valor de referencia el MSE obtenido de la etapa de Testing de los datos para seleccionar el resto de los indicadores.

Nodo Entrada	Nodo Oculto	MSE	RMSE	R	R2
1	10	0.010353118	0.101750271	0.65501224	0.429041034
2	13	0.006150878	0.078427536	0.820212968	0.672749313
3	10	0.00544876	0.073815719	0.835672884	0.69834917
4	10	0.005156232	0.071806907	0.843493293	0.711480936
5	8	0.004052034	0.063655588	0.879225427	0.773037352
6	9	0.003739816	0.061154036	0.887872317	0.788317252
7	6	0.003518596	0.059317752	0.897363236	0.805260777
8	8	0.00318704	0.056453874	0.909525093	0.827235895

9	2	0.003052257	0.055247231	0.911022195	0.82996144
10	14	0.002967269	0.054472646	0.913475033	0.834436637
11	2	0.002394524	0.048933872	0.930494169	0.865819399
12	13	0.002486728	0.049867108	0.927975697	0.861138893
13	4	0.002777861	0.052705418	0.919261196	0.845041146
14	2	0.002923258	0.05406716	0.91482799	0.836910251
15	2	0.003237042	0.056895012	0.909866675	0.827857366
16	6	0.003653748	0.060446241	0.907654812	0.823837258
17	2	0.003909674	0.062527386	0.887852939	0.788282842
18	6	0.003920756	0.062615943	0.895072735	0.8011552
19	2	0.004414264	0.066439929	0.873783282	0.763497224
20	2	0.004905026	0.070035891	0.853438751	0.728357702
21	3	0.005184934	0.072006487	0.849497519	0.721646034
22	2	0.004952081	0.070371026	0.855855335	0.732488354
23	2	0.005434225	0.073717195	0.832553125	0.693144706
24	8	0.005223432	0.072273318	0.855273886	0.73149342
25	2	0.005040194	0.070994321	0.859210709	0.738243043
26	8	0.004822432	0.069443735	0.860919277	0.741182002
27	3	0.005799166	0.076152252	0.81780405	0.668803464
28	2	0.004969063	0.070491582	0.852015462	0.725930347
29	6	0.005030179	0.070923753	0.847000903	0.71741053
30	8	0.007154856	0.084586379	0.771378713	0.595025119
31	2	0.007480895	0.086492169	0.749973944	0.562460917
32	8	0.007011344	0.083733769	0.775082987	0.600753636
33	4	0.006579388	0.081113429	0.812290282	0.659815502
34	3	0.006443648	0.080272334	0.802202193	0.643528359
35	2	0.007585465	0.087094574	0.766888508	0.588117984
36	2	0.006132597	0.078310902	0.81396369	0.662536889
37	2	0.006060329	0.077848112	0.822013691	0.675706508
38	3	0.006866163	0.082862314	0.78633977	0.618330234
39	12	0.008802723	0.09382283	0.721965294	0.521233885
40	6	0.008109927	0.090055133	0.729440964	0.53208412
41	3	0.007688736	0.087685438	0.746141187	0.556726671
42	10	0.008620334	0.092845754	0.715558111	0.512023411
43	4	0.008782879	0.093717016	0.715584481	0.512061149
44	8	0.009201363	0.095923735	0.699967404	0.489954367
45	2	0.008813656	0.093881075	0.705027764	0.497064148

46	2	0.008389243	0.091592813	0.722432425	0.521908609
47	2	0.008329799	0.091267731	0.723697102	0.523737495
48	3	0.008507369	0.092235402	0.718816362	0.516696962

Metricas					
Minimo	2	0.002394524	0.048933872	0.65501224	0.429041034
Maximo	14	0.010353118	0.101750271	0.930494169	0.865819399
Desviacion Estandar	3.6538	0.00210859	0.01413977	0.07485845	0.12113198
Promedio	5.2708	0.00573469	0.07442396	0.82295829	0.68274739

Tabla 9: LPSO - Lesionados Otras Comunas

Como se puede observar en los resultados de MSE la mejor configuración que se obtuvo para este set de datos fue de 11 nodos de entradas con 2 nodos ocultos. Lo cual nos indica que el modelo puede predecir con buenos resultados un desfase de 11 semanas. Los resultados nos demuestran que para este set de datos el modelo pudo predecir con un R^2 de 86 %. Como se puede observar en el cuadro de métricas el promedio de los 48 nodos ocultos nos dio un R^2 de 68 % el cual nos indica que para este set de datos el modelo predijo con buenos resultados para casi todos los nodos de entrada. Como se puede observar en la desviación estándar los valores son bastante bajos lo que nos demuestra que la variabilidad de los resultados no es tan amplia.

En la figura 17 se grafica los resultados obtenidos para la configuración de 11 nodos de entradas y 2 nodos de salida, donde se obtuvieron mejores métricas.

Como se puede apreciar en el gráfico de error relativo se puede observar, que una importante fracción de la información analizada es aceptable debido a que los valores residuales son bajos, con una banda de ajuste que van desde

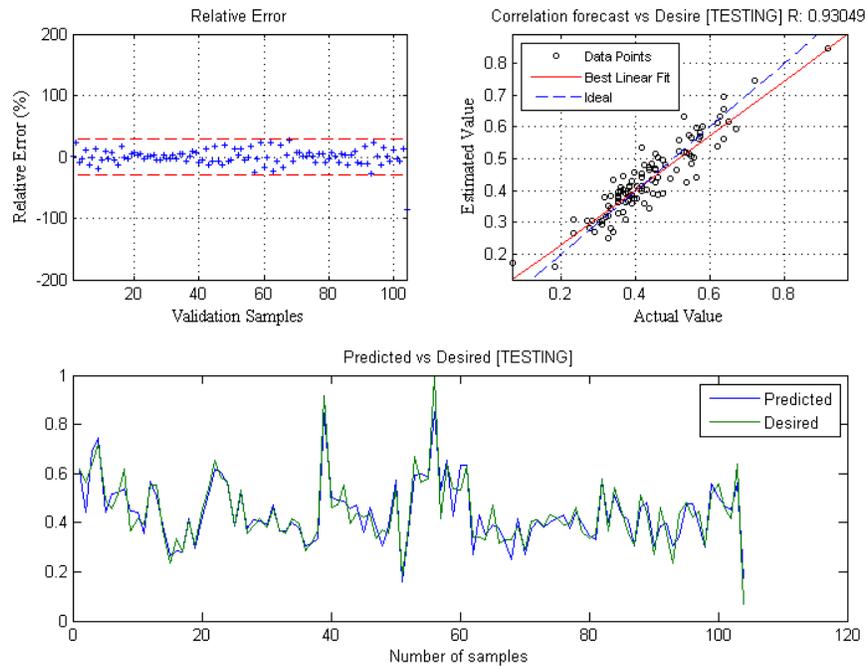


Figura 17: Grafico 11 nodos de entrada, 2 ocultos

−25 % a 25 %. En el gráfico de la correlación pronosticada podemos observar que la mayoría de los puntos se están agrupando en la cercanía de la recta ideal. Para el modelo la recta pronosticada se asemeja bastante a la recta ideal. Como se puede observar en el gráfico de la data pronosticada v/s deseada la gráfica representa una curva que se asemeja bastante a la data deseada a pesar de tener un R^2 que rodea el 86 %.

7.9. Consolidación de resultados obtenidos

Como se puede observar en tabla 10 los resultados de no difieren mucho entre todos los resultados obtenidos. Tenemos una varaiación de R^2 que fluctua entre el 86 % y el 89 % lo que se considera unos valores optimos. Podemos destacar que los mejores resultados se obtuvieron analizando los lesionados del Gran

Valparaiso con PSO y LPSO.

Analisis	Comuna	Tipo	Nodo Entrada	Nodo Oculto	MSE	RMSE	R	R^2
Muertos	Gran Valparaiso	PSO	16	7	0.00322	0.05680	0.93320	0.87086
Lesionados	Gran Valparaiso	PSO	12	3	0.00123	0.03515	0.94414	0.89141
Muertos	Otras Comunas	PSO	13	3	0.00398	0.06313	0.93341	0.87126
Lesionados	Otras Comunas	PSO	18	2	0.00246	0.04966	0.93652	0.87707
Muertos	Gran Valparaiso	LPSO	11	3	0.00343	0.05860	0.93324	0.87093
Lesionados	Gran Valparaiso	LPSO	13	3	0.00123	0.03516	0.94412	0.89136
Muertos	Otras Comunas	LPSO	11	12	0.00438	0.06625	0.93233	0.86924
Lesionados	Otras Comunas	LPSO	11	2	0.00239	0.04893	0.93049	0.86581

Tabla 10: Resumen de Resultados.

8. Conclusión

A través del desarrollo del presente proyecto se ha logrado entender el funcionamiento base de las redes neuronales artificiales tratadas, pudiéndose razonar sus principales ventajas y desventajas, lo que las hace una poderosa herramienta, la cual puede seguir evolucionando. También se pudo rescatar las diferencias entre las redes expuestas, en donde destaca como principal diferencia su función de activación.

Se logró una comprensión del algoritmo PSO y de algunas de sus variantes, con lo cual se logró dar cuenta de su simplicidad y eficiencia, razones que permiten justificar la elección de este algoritmo de computación evolutiva, como medio para el entrenamiento de las redes neuronales artificiales.

También se ha expuesto la problemática a resolver enmarcando los límites del estudio a realizar y las técnicas a utilizar, junto al análisis de los datos que se pretenden estudiar en el informe para obtener la relación entre los cambios de cantidades de los años.

Y además se ha propuesto un modelo de estimación de personas fallecidas para la V región de Chile. El cual busca a través de redes neuronales con algoritmo de optimización PSO y una variante para encontrar las mejores soluciones.

La información a analizar correspondió a los datos de los accidentes entre los años 2002 y 2012. Los cuales fueron separados en 2 grupos de comunas, para luego ser ordenados en grupos semanales de datos, cuales contienen la cantidad de personas fallecidas y lesionadas. Estos datos fueron analizados por la red neuronal con un modelo de desfase que representa la cantidad de semana que

puede predecir.

Al analizar los resultados nos damos cuenta que los mejores resultados (para todos los grupos de datos) se encontraron donde el número de nodos de entradas es alto, sin importar si los resultados fueron obtenidos de los fallecidos o lesionados. Lo que nos da por conclusión que el modelo obtiene los mejores resultados para predicciones con mayor número de desfase de la red. Con valores que superan un 85 % de Coeficiente de Determinación (Varianza Residual). Las métricas de rendimiento para todas las pruebas promediaron valores mayores de 68 %, para R^2 , lo que nos demuestra que la tendencia de buenos resultados fue bastante alta y eso se vio reforzado por los valores obtenidos en la desviación estándar. Los resultados también nos ayudan a concluir que no existe mucha diferencia entre las agrupaciones realizadas, ya que no hubo mayores diferencias entre los resultados obtenidos con fallecidos y lesionados ya sean del Gran Valparaíso o de el resto de las comunas en análisis.

Referencias

- [1] A. Abraham, H. Guo, and H. Liu. Swarm intelligence: Foundations. *Perspectives and Applications, Studies in Computational Intelligence (SCI)*, Springer-Verlag, 26:3–25, Noviembre 2006.
- [2] Ali Payidar Akgungor and Erdem Dogan. An application of modified smeed, adapted andreassen and artificial neural network accident models to three metropolitan cities of turkey. *Scientific Research and Essay*, pages 906–913, July 2009.
- [3] Felix Castro. *Ls-svm basado en optimización por enjambres de partículas para clasificación de accidentes de tránsito*. PhD thesis, Pontificia Universidad Católica de Valparaíso, Chile, 2010.
- [4] Miao M. Chong, Ajith Abraham, and Marcin Paprzycki. Traffic accident analysis using decision trees and neural networks. *CoRR*, 2004.
- [5] Estadísticas generales. CONASET, Agosto 2011.
- [6] F. Izaurieta and C. Saavedra. *Redes Neuronales Artificiales*. PhD thesis, Universidad de Concepción, Concepción, 2013.
- [7] J. Kennedy and R. Eberhart. Particle swarm optimization. *In Proceedings of the 1995 IEEE International Conference on Neural Networks, Piscataway, New Jersey*, 1995.

- [8] J. Kennedy, R. Eberhart, and Y. Shi. Swarm intelligence. *The Morgan Kaufmann Series in Artificial Intelligence*, Morgan Kaufmann, San Francisco, pages 3–80, April 2001.
- [9] José David Fierro Moya. *Clasificación de accidentes de tránsito en la zona de Valparaíso utilizando redes neuronales con algoritmos PSO*. PhD thesis, Pontificia Universidad Católica de Valparaíso, Chile, 2012.
- [10] L Mussone, A Ferrari, and M Oneta. An analysis of urban collisions using an artificial intelligence model. *Accident analysis and prevention*, 31(6):705–718, 1999.
- [11] Juan Zuñiga. Aplicación de redes bayesianas para identificar factores determinantes en accidentes de tránsito que afecten a las personas. *Technical report*, Universidad de Santiago de Chile, 2008.
- [12] P. Isasi Vinueza y I. Galván León. *Redes Neuronales Artificiales Un enfoque Práctico*. Pearson Educación, 2004.
- [13] World Health Organization. *Informe Sobre la Situación Mundial de la Seguridad Vial: Es Hora de Pasar a la Acción*. World Health Organization, 2009.
- [14] D. Poole, A. Mackworth, and R. Goebel. *Computational Intelligence: A Logical Approach*. Oxford University Press, New York, 1998.
- [15] A. J. Serrano, E. Soria, and J. D. Martín. *Redes Neuronales Artificiales*. PhD thesis, Universidad de Valencia, 2009.

- [16] So Young Sohn and Sung Ho Lee. Data fusion, ensemble and clustering to improve the classification accuracy for the severity of road traffic accidents in korea. *Safety Science*, 41(1):1–14, 2003.