

**CLASIFICACION DE ACCIDENTES DE TRANSITO UTILIZANDO REDES
NEURONALES CON ALGORITMO PSO****TRAFFIC ACCIDENT CLASSIFICATION WITH NEURAL NETWORKS WITH PSO
ALGORITHM****Jose Fierro Moya**Pontificia Universidad Católica de Valparaíso
Escuela de Ingeniería civil informática
e-mail : fierro.moya@gmail.com**Jose Fierro Moya**Pontificia Universidad Católica de Valparaíso
Escuela de Ingeniería de Transporte
e-mail : cmontt@ucv.cl**Nibaldo Rodriguez Agurto**Pontificia Universidad Católica de Valparaíso
Escuela de Ingeniería civil informática
e-mail : nibaldo.rodriquez@ucv.cl**RESUMEN**

La alta tasa de accidentes de tránsito en el mundo con consecuencias fatales hace necesario clasificarlos en forma aceptable. En este trabajo se estudiarán las redes neuronales artificiales y algoritmos evolutivos para lograr la clasificación del resultado del accidente, es decir, determinar si una persona resulta ileso o lesionado. Las redes neuronales artificiales entrenan sus pesos para lograr una buena clasificación, para lo cual se seleccionó el algoritmo evolutivo PSO. Se hicieron comparaciones de los resultados obtenidos, en base al promedio de distintas muestras con QPSO, con trabajos relacionados que no utilizan un promedio de muestras, sino que, el mejor para una muestra en particular, por lo cual, las diferencias podrían ser aún más elevadas con el modelo obtenido en este trabajo. Para la exactitud, el mejor resultado lo presenta LSSVM entrenado con PSO FIL, seguido de la red neuronal artificial tangente hiperbólica sigmoideal entrenada con QPSO y de redes bayesianas.

.Palabras claves: **data-mining, neural networks and traffic accidents**

ABSTRACT

The high rate of traffic accidents in the world with fatal consequences is necessary to classify them as acceptable. In this paper we study the neural networks and evolutionary algorithms to achieve the classification result of the accident, ie whether a person is uninjured or injured. Artificial neural networks trained their weights to achieve a good classification, which was selected for the evolutionary algorithm PSO. Comparisons of results obtained from the average of different samples with QPSO, related jobs that do not use an average of samples, but the best for a particular sample, therefore, differences could be even more high with the model obtained in this work. For accuracy, the best result has trained with PSO LSSVM IDF, followed by the artificial neural network trained with hyperbolic tangent sigmoideal QPSO and Bayesian networks.

KEYWORDS. First keyword. Second keyword. Last keyword.

1. Introducción

La novena causa de mortalidad mundial corresponde a las personas en las vías de tránsito, se estima que para el año 2030 escale hasta la quinta posición, OMS (2009). Los accidentes de tránsito forman parte de los 1,44 millones de adolescentes que mueren cada año, según informe de La UNICEF de Abril del 2012.

En Chile la situación no es menor, los siniestros de tránsito se han posicionado como una verdadera epidemia social, alcanzando cifras alarmantes y complejas de abordar. Según la Comisión Nacional de Seguridad de Tránsito (CONASET), solamente en 2009 se contabilizaron un total de 73.577 accidentes, siendo la colisión el tipo de siniestro más recurrente. La región de Valparaíso de Chile, analizada en este trabajo, es la segunda región con más siniestro a nivel nacional y el 2009 contribuyó con un total de 6.662 siniestros, en los cuales se perdieron 134 vidas humanas, CONASET (2009). En Chile los costos por accidentes de tránsito en el 2009 llegaron a un total de 357.462.613.856 pesos chilenos, esto considerando los costos sociales y privados. Estos costos consideran un total de 73.577 de siniestros de tránsito de los cuales 6662 pertenecen a la región de Valparaíso, es decir, aproximadamente un 11.8% de los siniestros de tránsito del País.

Una de las interrogantes que surge es como lograr una aceptable clasificación de los accidentes en estos siniestros, específicamente, para así poder conocer mejor las consecuencias de determinadas condiciones de tránsito para tomar conciencia y reducir los costos tanto sociales como privados. Dado lo anterior en este trabajo se estudiarán las redes neuronales artificiales y algoritmos evolutivos para lograr así una clasificación en cuanto al tipo de accidente, es decir, determinar si una persona resulta ileso o lesionada en un accidente de tránsito, dado los atributos de entrada de la red, que son por ejemplo: comuna; edad; causa; conductor; pasajero y peatón choque; entre otros.

Actualmente si bien hay investigaciones en accidentes de tránsito mediante diversas técnicas (redes Bayesianas, MLP y SVM entre otras), el uso de redes neuronales artificiales en clasificación de accidentes de tránsito es reducido y su enfoque dependerá de cómo el investigador desee construir el modelo, es por ello que esta investigación se centra en el uso de redes neuronales artificiales para la clasificación de accidentes de tránsito en Valparaíso Chile. Por lo anterior se desarrollará un modelo de clasificación del grado de severidad de las lesiones con que resultan las personas en accidentes de tránsito en la zona de Valparaíso de Chile utilizando redes neuronales con algoritmos de optimización por enjambre de partículas.

Las redes neuronales artificiales requieren entrenar sus pesos para lograr una buena clasificación, es por ello que se ha seleccionado el algoritmo evolutivo PSO para dicha tarea. Si bien existen otros métodos como Back Propagation, se eligió PSO por su simplicidad y eficiencia (S. Mohaghegi et al 2011).

Las funciones de activación para los nodos de la capa oculta fueron la función sigmoideal o la función tangente hiperbólica. Las funciones de activación para los nodos de la capa de salida fueron la función sigmoideal o la función tangente hiperbólica en caso de una neurona de salida, la cual representa el estado en que quedo la persona en el accidente de tránsito, ileso o lesionado. Se realizó un entrenamiento mediante variantes de PSO, entre ellas se encuentran: LDWPSO, QPSO, LDWQPSO. La cantidad de partículas utilizadas para las variantes de PSO fue de 100. En cuanto a la función de desempeño utilizada por estas variantes de PSO, estas fueron: error basado en la exactitud y entropía cruzada.

Los resultados obtenidos son comparados entre las diferentes técnicas utilizadas mediante tablas comparativas que facilitan su comprensión.

2. Metodología

Para llegar a la obtención de los modelos se han de seguir los siguientes pasos, que se muestran en la figura 2.1.

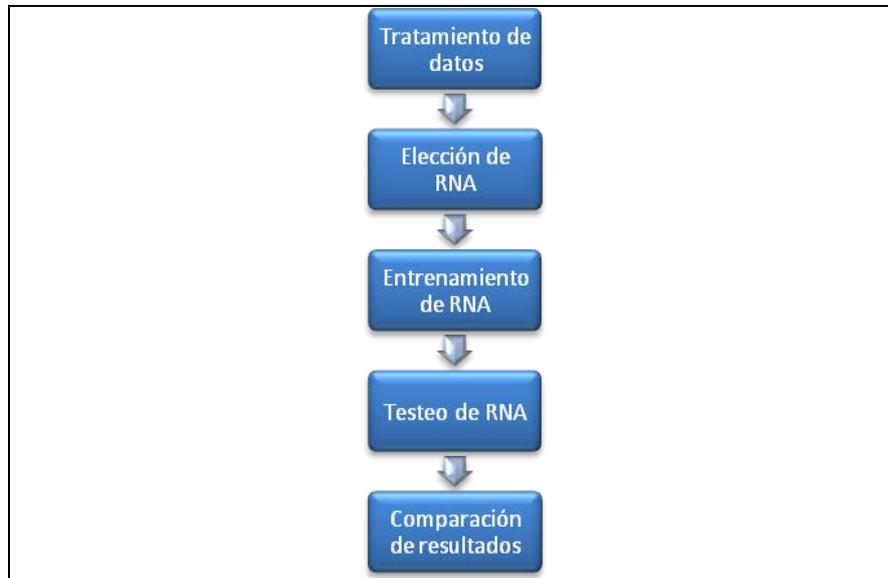


Figura 2.1 Pasos para el desarrollo del modelo

1. **Tratamiento de datos:** se trataron los datos obtenidos de la CONASET (Comisión Nacional de Seguridad de Tránsito) que datan desde el año 2003 hasta el año 2009 en la región de Valparaíso de Chile, para ello se realizó un pre-proceso de los datos en el cual se eliminaron registros que presentaran datos redundantes o inconsistencias para luego normalizarles facilitando la labor posterior. Se utilizaron 73.577 accidentes de tránsito, que corresponden a los de ellos se utilizaron diversas muestras (80% para entrenamiento y 20% para prueba).
2. **Elección de red neuronal artificial y algoritmo de entrenamiento:** en esta etapa se seleccionan los tipos de redes usados para realizar los diversos modelos, para esta investigación se utilizará MLP, en cada una de ellas se probarán diversas funciones como se verá más adelante. Esta etapa considera:
 - a. **Selección de algoritmo de entrenamiento:** el algoritmo de entrenamiento a utilizar será LDWPSO, QPSO y/o LDWQPSO.
 - b. **Elección de topologías:** los nodos de entrada de la red son 12, Comuna Urbano/Rural; estado Atmosférico; hora; causa; tipo calzada; estado calzada ;condición calzada; tipo de accidente; calidad; sexo; edad y resultado. Los nodos de la capa oculta se elegirán según LDWPSO, QPSO y/o LDWQPSO. La capa de salida presenta un nodo de salida.
3. **Entrenamiento de la red:** la red se entrena mediante LDWPSO, QPSO y/o LDWQPSO de manera de obtener los pesos de la red, para ello se utilizará un determinado número de datos como mostrará más adelante, por ejemplo un 80% de los datos.

4. Testeo de la red: se obtienen los resultados una vez ya entrenada la red, para ello se utilizará un determinado número de datos, los cuales no hayan sido utilizados para el entrenamiento, por ejemplo un 20% de los datos
5. Comparación de resultados: en la etapa final se compararán los resultados obtenidos por los diversos modelos mediante métricas especificadas más adelante.

2.1 Tratamiento de datos

Para el tratamiento de los datos, primero estos son normalizados lo que permite por lo general obtener mejores resultados, esto al acotar el rango en el cual se encuentran los datos, es por ello que los datos mencionados anteriormente fueron normalizados mediante la siguiente fórmula aplicada a cada atributo por separado:

$$z = \text{factor} * [(2 * x - (\max(x) + \min(x))) / (\max(x) - \min(x))]$$

Donde el **factor** tiene un valor de 0.8 o 0.9, $\max(x)$ representa el máximo valor existente en un atributo, $\min(x)$ representa el mínimo valor existente en un atributo.

2.2 Red Neuronal Base

La red neuronal base, es el PERCEPTRON multicapa (MLP, según sus siglas en inglés Multi Layer PERCEPTRON) o red multicapa con conexiones hacia adelante, el cual se muestra en la siguiente figura, surge como consecuencia de las limitantes presentadas por el PERCEPTRON simple en lo relacionado al problema de separabilidad no lineal.

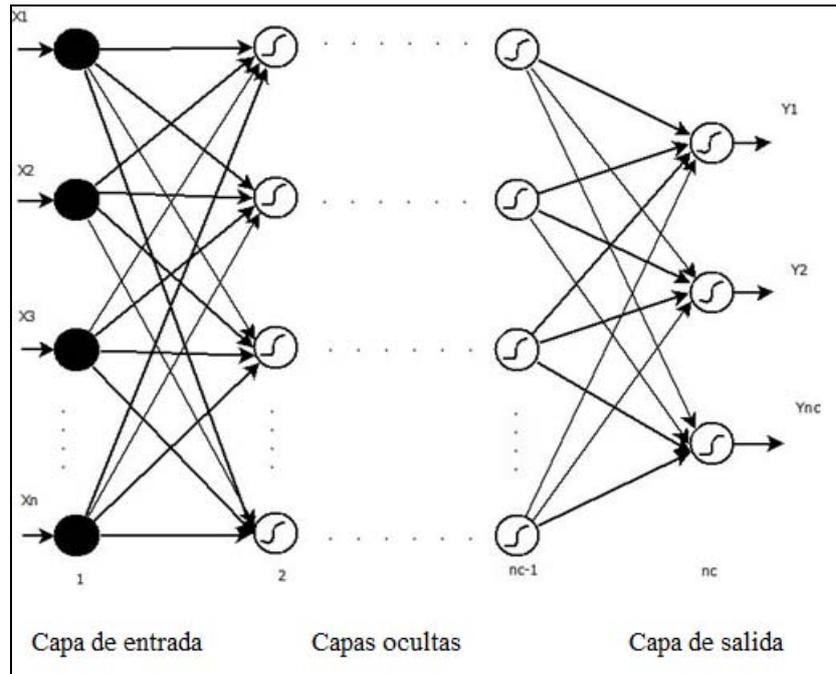


Figura 2.2 Red neuronal MLP

Como se puede apreciar en la figura 2.2, este tipo de redes consta de una capa de entrada por donde se ingresan los datos, capas ocultas (una o más) y una capa de salida, el resultado estará determinado por los datos de entrada, los pesos entre las neuronas artificiales y las funciones de activación de las capas.

Las funciones de activación más utilizadas son:

La función sigmoideal, cuya imagen está en el rango [0,1]

$$f_1 = \frac{1}{1 - e^{-x}} \quad 2.1$$

La función tangente hiperbólica, cuya imagen está en el rango [-1,1]

$$f_2 = \frac{1 - e^{-x}}{1 + e^{-x}} \quad 2.2$$

2.3 Entrenamiento de la red neuronal

La red se entrenó con el método Particle Swarm Optimization (PSO) Este algoritmo simula el vuelo de una bandada de pájaros en busca de alimento.

Cada partícula tiene una velocidad, posición, la mejor posición personal encontrada hasta el momento. El enjambre posee una mejor posición de todo el enjambre. Es así que el movimiento de cada partícula estará determinado por la posición particular de la partícula y la del enjambre. A continuación se explica cómo funciona:

1.- Evaluar la función fitness para cada partícula, que significa que posición tiene mejor resultado para la red, en este trabajo la función de fitness indica el grado de exactitud con que se clasificó el estado de la persona correctamente o como el mínimo error de clasificación que entrega la red. Como lesionada o ilesa

2.- Comparar la función fitness actual de cada partícula con su mejor personal (P). Si X_i es mayor que P_i , entonces P_i toma el valor X_i . Si X_i es mayor que G_i , entonces G_i toma el valor X_i .

3.- Ajustar la posición actual y la velocidad de cada partícula según las siguientes ecuaciones, donde w es el factor de inercia que establece en qué medida la nueva velocidad estará influenciada por la anterior, C_1 establece en qué medida la nueva velocidad estará influenciada por el comportamiento personal de la partícula, C_2 establece en qué medida la nueva velocidad estará influenciada por el comportamiento grupal del enjambre, p es la mejor posición personal y g es la mejor posición grupal:

$$\begin{aligned} v_{ij}^{t+1} &= w \times v_{ij}^t + C_1 \times rand() \times (p_{ij}^t - x_{ij}^t) + C_2 \times rand() \times (g_j^t - x_{ij}^t) \\ x_{ij}^{t+1} &= x_{ij}^t + v_{ij}^{t+1} \end{aligned} \quad 2.3$$

Linearly Decreasing Weight Particle Swarm Optimization o LDWPSO

Fue propuesto por Shi y Eberhart en 1998, en él se consideran las ecuaciones anteriores de manera que el factor de inercia va disminuyendo su valor conforme aumenta el número de iteraciones. Esto mejora el desempeño con respecto al PSO original al equilibrar las habilidades de búsquedas globales y locales. Para lograr esto el peso de inercia del LDWPSO decrece linealmente de 0.9 a 0.4, su valor en cada iteración estará dado por la siguiente ecuación:

$$\omega_{LDW} = (\omega_{\max} - \omega_{\min}) \times \frac{\text{iteración } \omega_{\max} - \text{iteración } \omega_i}{\text{iteración } \omega_{\max}} + \omega_{\min} \quad 2.4$$

Donde ω_{\max} es 0.9, ω_{\min} es 0.4 y iteración max es el máximo de iteraciones permitidas.

Métricas de Desempeño

La entropía cruzada se utilizó en este trabajo para determinar los pesos de la red neuronal artificial mediante LDWPSO.

La siguiente ecuación considera una sola neurona de salida:

$$E = -\frac{1}{N} \times \sum_{i=1}^N (d_i(n) \times \ln(y_i(n)) + (1 - d_i(n)) \times \ln(1 - y_i(n))) \quad 2.5$$

La siguiente ecuación considera dos neuronas de salida:

$$E = -\frac{1}{N} \times \sum_{i=1}^N \left[-\sum_{l=1}^J d_l(n) \times y_l(x(n), \theta(l)) + \log \sum_{l=1}^J e^{y_l(x(n), \theta(l))} \right] \quad 2.6$$

Para determinar los resultados de la red se requiere de métricas acordes a un modelo de clasificación para determinar el rendimiento del modelo, para lo cual es preciso conocer primero los siguientes términos:

Verdaderos Positivos o VP: también conocidos como éxitos, corresponde al número de personas clasificadas por el modelo como lesionadas y resultan realmente lesionadas.

Falso Positivo o FP: también conocidos como falsas alarmas, corresponde al número de personas clasificadas por el modelo como lesionadas, pero que resultan realmente ilesas.

Verdaderos Negativos o VN: también conocidos como rechazos correctos, corresponde a las personas clasificadas por el modelo como ilesas y resultan realmente ilesas.

Falsos Negativos o FN: corresponde al número de personas clasificadas por el modelo como ilesas, pero que resultan realmente lesionadas.

Esto se puede visualizar en la siguiente tabla 2.1, en donde se señalan los tipos de error.

Tabla 2.1 Matriz de Confusión

		Resultado Real	
		Lesionadas	Ilesas
Resultado obtenido por el modelo	Lesionadas	VP	FP, error tipo II
	Ilesas	FN, error tipo I	VN

Teniendo claro lo anterior es que se pueden establecer las siguientes métricas para determinar el rendimiento del modelo de clasificación:

- Sensibilidad: indica la probabilidad de clasificar a un individuo como lesionado, es decir el porcentaje de son clasificadas efectivamente como lesionadas. La sensibilidad está definida por la siguiente ecuación:

$$\text{Sensibilidad} = \left(\frac{VP}{VP + FN} \right)$$

- Especificidad: indica la probabilidad de que un individuo ileso sea realmente clasificado como ileso. La especificidad está definida por la siguiente ecuación:

$$\text{Especificidad} = \left(\frac{VN}{VN + FP} \right)$$

- Exactitud: indica el porcentaje de personas bien clasificadas, tanto lesionadas como ilesas. La exactitud está definida por la siguiente ecuación:

$$\text{Exactitud} = \left(\frac{VP + VN}{VP + VN + FP + FN} \right)$$

3. Análisis de datos.

La tabla 3.1 muestra el modelo que presenta la mejor exactitud promediada de las muestras, este considera una red neuronal con 12 nodos de entrada, 4 nodos de capa oculta y 1 de salida, siendo sus funciones de activación sigmoideal para la capa oculta y tangente hiperbólica para la capa de salida.

Tabla 3.1 resultados con mejor exactitud

Muestras	Tangente hiperbólica-Sigmoideal		
	LDWPSO	QPSO	LDWQPSO
1000	0,8150	0,8350	0,7950
3000	0,8017	0,7950	0,7717
10000	0,8100	0,80,65	0,8135
20000	0,8070	0,80,77	0,7887
Promedio de las muestras	0,8084	0,8111	0,7922

En la tabla 3.2, se presenta el modelo que entregó mejores resultados respecto a la exactitud, este considera una muestra de 3000 datos y en una red neuronal sigmoideal-sigmoideal entrenada mediante LDWPSO con 9 nodos ocultos y un nodo de salida.

Tabla 3.2 Resultado de Modelo con una muestra de 3000 datos

PSO	Funciones	Numero de Nodos	Exactitud	Especificidad	Sensibilidad

		ocultos			
LDWPSO	Sigmoidal-sigmoidal	9	0,8233	0,9483	0,7065

También se utilizó un modelo utilizando 3 neuronas de entrada las cuales representan los parámetros calidad, causa y edad. Las neuronas de la capa oculta utilizaron la función sigmoidal como función de activación. Una neurona de salida que utilizó la función sigmoidal como función de activación.

La muestra fue de 3000 datos de los cuales el 80% se utilizó para determinar el número de neuronas ocultas.

El algoritmo para entrenar la red y determinar los mejores pesos tanto entre la capa de entrada y la oculta, así como entre la capa oculta y la de salida fue LDWPSO con un $C1=C2=1.45$, un número de partículas de 100, 300 iteraciones un $w_{max}=1$ y un $w_{min}=0.5$. La función de *fitness* fue entropía cruzada.

Se realizaron pruebas que consideraban desde 1 hasta 100 nodos ocultos repitiendo 5 veces el experimento y sacando el promedio para cada nodo oculto, es decir, se calculó el desempeño 5 veces para 1 nodo oculto y se calculó su promedio, así también para todos los demás y se determinó que cantidad de nodos ocultos ofrecía un mejor desempeño de la red, la entropía cruzada más baja.

Si se comparan los resultados promediados de las muestras que presentaron mejor exactitud porcentual para comparar el resultado del presente trabajo con los de Castro, 2010 y el realizado en Montt et al 2009, en este último se utilizó redes bayesianas y como métricas la exactitud, sensibilidad y especificidad, esto se puede apreciar la siguiente tabla 3.3.

Tabla 3.3 LS-SVM PSO vs Redes Bayesianas vs RNA QPSO

	LS-SVM PSO F.I.L	Redes Bayesianas	Red neuronal artificial Tangente Hiperbólica Sigmoidal QPSO
Sensibilidad	78,12%	71%	69,53%
Especificidad	87,01%	87%	93,55%
Exactitud	82,04%	78%	81,11%

En la tabla 3.3 queda claro que el modelo propuesto supera en más de tres puntos al de redes bayesianas, mas no así en comparación al modelo LS-SVM PSO F.I.L, el cual saco casi un punto más de exactitud.

Es importante señalar que estas comparaciones son realizadas en base al promedio de distintas muestras de los resultados obtenidos con QPSO, no así en los trabajos relacionados que no utilizan un promedio de varias muestras, sino más bien, el mejor resultado para una muestra en particular. Por ende, las diferencias podrían ser aún más elevadas con el modelo obtenido en este trabajo de investigación.

Como se puede observar en la tabla 3.3, respecto a sensibilidad el modelo LS-SVM entrenado con PSO F.I.L presenta los mejores resultados seguidos de Redes Bayesianas, siendo la red neuronal artificial tangente hiperbólica sigmoidal entrenada con QPSO la que presenta el resultado más bajo. Respecto a la especificidad, el mejor resultado lo presenta la red neuronal artificial tangente hiperbólica sigmoidal entrenada con QPSO, seguida de LS-SVM entrenado con PSO F.I.L y seguido de las redes bayesianas. En cuanto a Exactitud, el mejor resultado lo

presenta LS-SVM entrenado con PSO F.I.L, seguido de la red neuronal artificial tangente hiperbólica sigmoidal entrenada con QPSO y seguida de redes bayesianas.

Si analizamos los mejores modelos desde el punto de vista de la exactitud (ver tabla 3.4) respecto a la ganancia obtenida, se puede decir que:

- El modelo que presenta ganancia positiva en todas las comparaciones es LS-SVM entrenado con PSO F.I.L.
- La red neuronal artificial Tangente Hiperbólica Sigmoidal entrenada con QPSO presenta ganancia positiva respecto las redes bayesianas, no así respecto a LS-SVM entrenado con PSO F.I.L.
- Las redes Bayesianas presentan ganancia negativa frente a los otros modelos planteados en la tabla 3.4.

Tabla 3.1 Ganancia de los resultados promedio de la evaluación de los modelos de la tabla

A \ B	LS-SVM PSO F.I.L	Redes Bayesianas	Red neuronal artificial Tangente Hiperbólica Sigmoidal QPSO
LS-SVM PSO F.I.L	0%	4,9244%	1,1336%
Redes Bayesianas	-5,1795%	0%	-3,9851%
Red neuronal artificial Tangente Hiperbólica Sigmoidal QPSO	-1,1466%	3,8343%	0%

4. Conclusión

A través del desarrollo del presente proyecto se ha logrado entender el funcionamiento base de las redes neuronales artificiales tratadas, las cuales son una herramienta que puede seguir evolucionando. También se pudo obtener las diferencias con 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 evolutivo 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.

Las técnicas propuestas demostraron ser robustas y tener buenos resultados los que fueron contrastados y comparados con otros trabajos relacionados.

Comparando los resultados de es trabajo con el de otros trabajos, ,la mejor exactitud, es decir el modelo que logró la mejor clasificación, se obtuvo en Castro,2011 seguido por “Clasificación de Accidentes de Tránsito en la Zona de Valparaíso Utilizando Redes Neuronales con Algoritmos PSO” y por último “Estructura Bayesianas en Accidentes de Tránsito”.

Si bien el trabajo realizado presento diversas técnicas enfocadas en las redes neuronales artificiales, se pueden abordar muchas más, así como mejorar la forma de elegir el número de nodos de entrada y ocultos. Es así como el tema elegido permite ahondar más en futuros trabajos permitiendo la mejora de la técnica de clasificación, ya sea con diversas técnicas y/o métodos para entrenar estas.

Si la red se utiliza con tres entradas se pudo observar que se obtienen mejores resultados, y además se puede obtener de un análisis estadístico posterior, el porcentaje de ilesos o lesionados

en un tramo de edad, para una causa y un tipo de accidente, es decir se podría decir por ejemplo que las personas entre los 15-25 años en un porcentaje determinado son lesionadas por no respetar las señales de tránsito, por lo cual se produce una colisión. Lo que se realizara en investigaciones futura.

Como investigación futura es posible investigar técnicas que permitan determinar la mejor combinación de atributos de entrada, evaluando no solo la correlación lineal si no también la no lineal, de este modo se podría determinar los atributos de entrada no correlacionados, así también investigar otras técnicas de entrenamiento de redes menos estocásticas.

5. Bibliografía

- «CONASET, COMISIÓN NACIONAL DE SEGURIDAD DE TRÁNSITO,» 2010. [En línea]. Available: http://www.conaset.cl/portal/portal/default/estadisticas_generales.
- S. Mohaghegi, Y. del Valle, G. Venayagamoorthy y R. Harley, «COMPARISON OF PSO AND BACKPROPAGATION FOR TRAINING RBF NEURAL NETWORKS FOR IDENTIFICATION OF A POWER SYSTEM WITH STATCOM,» /.(2011) [En línea]. Available: <http://web.mst.edu/~ganeshv>
- F. Castro, 2010. LS-SVM basado en Optimización por Enjambres de Partículas para Clasificación de Accidentes de Tránsito, Valparaíso,
- C. Montt, A. Zúñiga y M. Chacón, , 2009 Estructura Bayesianas en Accidentes de Tránsito.
- Y. Shi y R. Eberhart, 1998«Parameter selection in particle swarm optimization: in Evolutionary Programming VII, Lecture Notes in Computer Science, 1447,» Berlin,.
- Catro, Montt, Rodriguez (2011) SVM “Basado en optimización por enjambres de Abejas” actas del IX Congreso Chileno de Investigación Operativa, OPTIMA 2011,, Temuco – Chile