CAPITULO 7. PRUEBAS Y RESULTADOS.

En este capítulo se describen las pruebas realizadas para comprobar el desempeño del sistema, obteniendo resultados cuantitativos que permiten apreciar la eficiencia del sistema.

7.1 PRUEBAS.

En primer lugar se debe resaltar el hecho de que las redes neuronales fueron puestas a prueba con datos desconocidos o nuevos, por lo que la función de transferencia de las neuronas de salida, no nos entregó exactamente los valores de uno o cero sino que nos entregó una aproximación de estos valores, es decir que nos entrego valores decimales entre cero y uno. Debido a que las cinco neuronas de salida presentaron distintos valores decimales, se paso la salida de las redes neuronales por una función de competencia, la cual, comparaba las cinco salidas y tomaba la de mayor valor, poniendo un uno en esa posición y cero en las demás posiciones, de manera que si a la salida de la red neuronal teníamos el vector [.1, .8, .2, .5, .3] a la salida de la función de competencia tendríamos el vector [0,1,0,0,0] indicando que el dato a la entrada de la red neuronal pertenece a la clasificación designada para la segunda neurona, la cual corresponde a un ataque de INYECCION DE COMANDOS, esto se puede apreciar claramente en la siguiente figura:

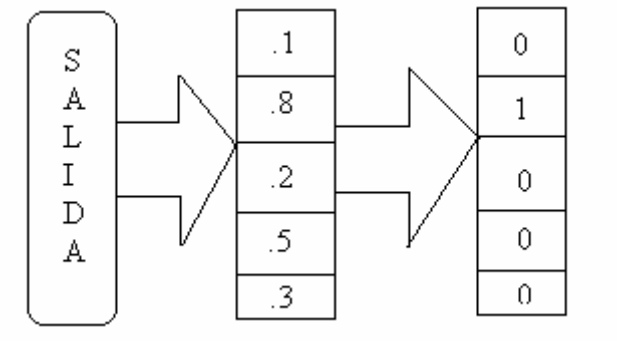

Figura 7.1 Función de Comptencia.

Es así como fuimos capaces de evaluar las redes neuronales al compararlas directamente con los destinos. Por otro lado es importante mencionar que en un principio el tamaño de los datos fue incrementado al hacer uso de la ventana deslizante, es decir, que en un principio contábamos con 773 cadenas de caracteres de tamaño variable, y al usar una ventana fija de 8 terminamos usando 2127 cadenas binarias de tamaño fijo representadas por una matriz de 2127 por 64. Es decir que el número de los destinos también fue incrementado de ser 773 destinos a ser 2127 destinos, sin embargo el interés era evaluar correctamente las 773 cadenas originales.

Normalmente una cadena de longitud N que hiciera referencia a un ataque del tipo X se convirtió a (N-7) cadenas de tamaño fijo igual a 8, por medio de la ventana deslizante. Por lo que los (N-7) destinos generados en teoría debieron haber arrojado el mismo resultado ya que hacían referencia a la misma cadena del tipo X. Sin embargo, debido a que la red neuronal entrega resultados aproximados, en la mayoría de los casos los destinos arrojados fueron distintos para las (N-7) salidas. Por lo que al evaluar la red, el destino que se tomo como válido para cada cadena, fue el promedio de los (N-7) resultados obtenidos para dicha cadena. Como ejemplo si la cadena [10, 23, 45, 68, 76, 102, 33, 46,12, 21] representara un ataque del tipo XSS, el sistema de detección de intrusos tomaría esta cadena de tamaño 10 y la dividiría en tres (10-7= 3) cadenas de tamaño 8 que serían las siguientes:

[10, 23, 45, 68, 76, 102, 33, 46]

[23, 45, 68, 76, 102, 33, 46,12]

[45, 68, 76, 102, 33, 46,12, 21]

Una vez divididas las cadenas, y convertidas a formato binario, serían ingresadas a la red neuronal por lo que en teoría todas deberían arrojar la siguiente salida [0 ,0 ,0 ,1 ,0], sin embargo es posible que las salidas hallan sido las siguientes:

Salida de Cadena 1 = [0, 1, 0, 0, 0]

Salida de Cadena 2= [0, 0, 0, 1, 0]

Salida de Cadena 3= [0, 0, 0, 1, 0]

Si nos damos cuenta, la primera fracción de la cadena arrojo un resultado distinto a las otras 2 cadenas, no obstante, el programa diseñado ha tomado como valida la salida [0, 0, 0, 1, 0] debido a que ha sido la salida que más se ha repetido.

El parámetro a evaluar en cada una de las redes neuronales diseñadas en el apartado anterior, es el porcentaje de identificación de ataque, y el porcentaje de detección de ataque. Para esto, se debe definir el porcentaje de identificación como el porcentaje con el cual la red acertó a expresar que un dato correspondía a una de las cinco posibles clasificaciones y resultó correcto. Es decir que cuando hablamos de la capacidad de la red para identificar el ataque nos referimos a la frecuencia con que los destinos fueron exactamente igual a las salidas entregadas por la red. Por otro lado, el porcentaje de detección únicamente se interesa por llevar el porcentaje con el cual la red acertó a identificar si un dato era un ataque o una acción normal, sin importarnos si el ataque fue correctamente identificado en la clasificación mencionada en el capítulo 2.4. La Figura 7.2 nos muestra una salida que puede ser a la vez desacierto de identificación y acierto de detección.

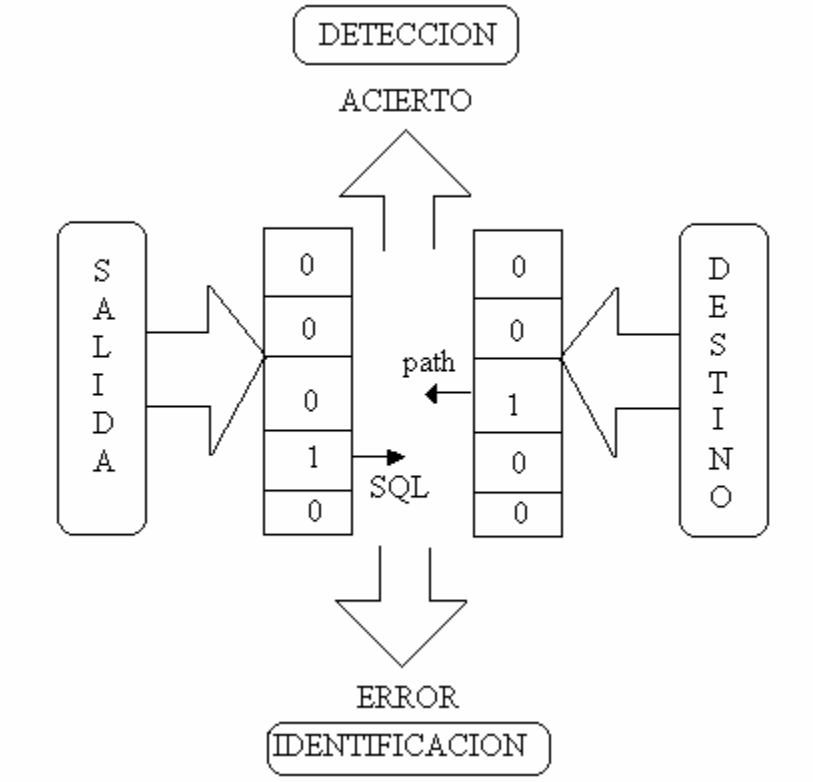

Figura 7.2 Identificación y Detección.

Como se puede inferir, el porcentaje de detección de la red neuronal deberá ser mayor al porcentaje de identificación ya que el primero no requiere tanta precisión como el segundo.

7.2 RESULTADOS PRACTICOS.

Al poner a prueba las distintas redes neuronales se observaron los siguientes resultados prácticos:

Tabla 7.1 Resultados Prácticos.

	MLP 64X15X15X5(*RBP*)	MLP 64X20X20X5 (*CG*)	ELMAN64X30X30X5(*RBP*)
% Identificación	84.85	84.42	87.01
% Detección	88.74	89.18	**90.91**

Como se puede observar, el porcentaje de detección fue más alto que el de identificación y se nota claramente que la red Elman generalizó mejor sobre los datos que las otras redes. La ventaja de la red Elman es que maneja una memoria de contexto que le permite tener seguimiento de los datos para hallar patrones de secuencia entre ellos. El proyecto desarrollado por [TOR03] maneja pruebas distintas ya que al implementar el sistema en forma práctica pudo probar su sistema con ataques reales. En nuestro caso se ha probado a la red neuronal con el 30% de los datos con los que contamos por lo que los resultados de estas pruebas tienen un significado distinto al ser únicamente pruebas de diseño y no de implementación.

Por otro lado, estos porcentajes pueden variar si la red se entrena con diferentes condiciones iniciales, y muchas veces se tiene que esperar a hallar las condiciones iniciales correctas para obtener los mejores resultados, de manera que no siempre vamos a obtener el mismo resultado. No obstante una vez que la red presente los porcentajes más altos, como es el caso, se puede concluir el entrenamiento y comenzar a utilizar la red como un sistema detector de intrusos.

Cabe resaltar la eficiencia de estos resultados al hacer notar que la red ha sido capaz de detectar hasta un 90% de los ataques en datos que fueron completamente desconocidos para la red. Es decir, que ha sido probada la capacidad de la red para generalizar en datos nuevos por lo que la obtención de estos resultados sugiere que la aplicación de las redes neuronales al campo de la detección de intrusos en la capa de aplicación es viable si se cuenta con los datos de entrenamiento suficientes para este propósito.

Es importante por lo tanto, que una vez que esta red sea implementada como un sistema de detección de intrusos, se este actualizando constantemente con nuevos ataques

adaptando los pesos de la red para que se tenga conocimiento de nuevos ataques, y capacidad para generalizar sobre los mismos.