Universidad Simón Bolívar

Departamento de Computación y T.I.

Inteligencia Artificial II

Septiembre-Diciembre 2014

Estudiantes: Carnets:

Wilmer Bandres 10-10055

Juan Escalante 10-10227

**PROYECTO 1 - PARTE 2: REDES MULTICAPAS**

Para esta segunda entrega se implementó el algoritmo de feedforward con backpropagation en una red de múltiples capas que consta de tres capas (la capa del input, la capa intermedia y la capa del output) y haciendo uso de funciones sigmoidales que son usadas ampliamente en todo lo que se refiere al tema de las redes neuronales.

Inicialmente se entrenó la red para un ejercicio particular, que consistía en cierta cantidad de puntos con el objetivo de poder clasificar si un punto está en el interior o en exterior de un círculo. El entrenamiento se realizó con casos tanto generados de forma aleatoria como ya proporcionados para el ejercicio y las pruebas se realizaron sobre un conjunto de 10.000 puntos generados aleatoriamente con resultados muy buenos, dando un porcentaje de error muy bajo.

Posteriormente se entrenó la red para la tarea de clasificar 150 plantas y separarlas en dos grupos, aquellas plantas que son Iris Setosa y aquellas que no lo son. La red se entrenó con un conjunto del 50%, 60%, 70%, 80% y 90% de las 150 plantas y las pruebas se realizaron sobre el 100% de los datos.

Finalmente la red fue entrenada para clasificar las mismas 150 plantas del ejercicio anterior, pero esta vez en tres grupos: Iris Setosa, Iris Versicolor e Iris Virginica. El entrenamiento y las pruebas de este ejercicio son similares también a las del ejercicio anterior.

* **Detalles de implementación**

1. Se realizó una clase llamada mi\_vector para representar el vector de pesos de cada neurona y facilitar las operaciones de feedforward en la red.
2. Existen 3 implementaciones de redes, que en realidad son las mismas solo que se diferencian en el tipo de problema que tratan. En el README del proyecto dice como correr cada uno de los ejercicios, además identifica a qué ejecutable corresponde cada ejercicio.
3. Para la realización de las gráficas se utilizó gnuplot.
4. Para la creación de gráficas y para la generación todos los resultados de los casos de pruebas, ya que eran tantos, y había que variar el número de neuronas de 4 a 10, se implementaron scripts que realizan todo de forma rápida.
5. El máximo de iteraciones del aprendizaje es de 50000 o puede ser menos en caso de que el error cuadrático medio converja a un número muy pequeño (0.001).
6. En el caso del primer ejercicio de las iris, que consistía en clasificar las plantas de forma binaria (es setosa o no), se decidió colocar, a la hora de parsear la entrada, valores de 1’s (si es setosa) y 0’s (si no es setosa) en lugar de 1’s y -1’s ya que se obtuvieron resultados mejores. Similarmente en el segundo ejercicio de las plantas. En el caso del ejercicio del círculo también se obtuvieron mejores resultados con estos valores.
7. La función sigmoidal utilizada fue la siguiente: sigmoidal(x) = 1/(1+exp(-x)).

* **Detalles de experimentación**

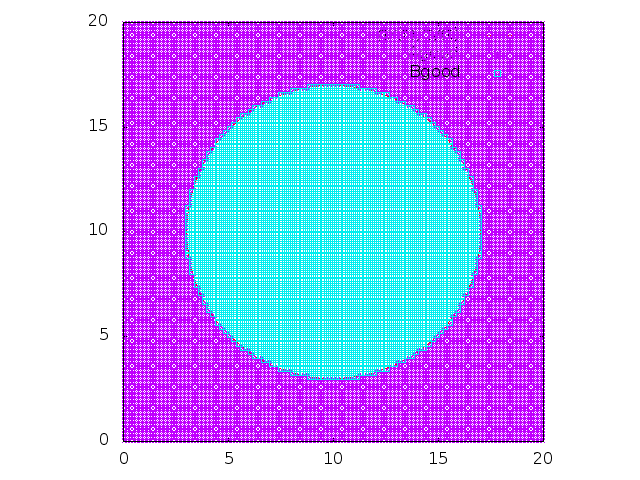
1. Se dejan las iteraciones hasta 50.000 para poder contemplar en el gráfico de iteraciones-errores aquellos puntos donde el error esta bajando y de repente aumenta pero luego vuelve a decaer. Hecho importante que nos ayuda visualmente a observar cuando realmente debemos parar de iterar ya que no necesariamente es la primera vez que el error aumente un poco.
2. La tasa utilizada para todos los experimentos es de 0.1, ya que no se consideró como un parámetro de tanta importancia como sí lo fue el número de neuronas de la capa intermedia o el porcentaje del archivo de entrenamiento como es el caso del ejercicio de las iris. El cambio de la tasa simplemente traía cambios mínimos y por eso se decidió dejar una única tasa para todos los experimentos.
3. Todos los resultados de los experimentos se encuentran bajo la carpeta experimentos. Y están ubicados bajo la carpeta llamada "círculo” (aquellos experimentos que se refieren al primer ejercicio), la carpeta "setosa" (aquellos experimentos referentes a la división de las plantas setosas de las demás) y la carpeta "divisor\_de\_iris" (aquellos experimentos del último ejercicio).
4. Cada archivo de salida de algún experimento se encuentra bajo la extensión .out. Mientras que cada archivo de salida iteración/error se encuentra bajo la extensión .txt. Y cada gráfico se encuentra con la extensión .png.
5. Para correr cada tipo de experimento solo es necesario leer el README.

* **Resultados**

Para cada uno de los ejercicios se mostrarán los resultados más relevantes.

* + En el problema de los puntos, se realizaron múltiples experimentos, variando tanto el número de neuronas de la capa intermedia como el conjunto de entrenamiento. El parámetro más importante que hacía que la capa clasificará de mejor forma los puntos fue el número de neuronas de la capa intermedia, al parecer existe una tendencia en donde dado un mismo conjunto de entrenamiento, mientras mayor sea la cantidad de neuronas en la capa intermedia, el porcentaje de error de prueba disminuye y esto se evidencia a continuación:

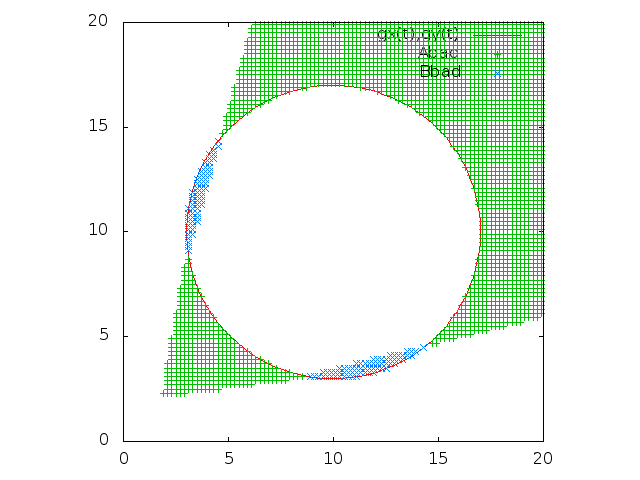
Primero que nada, en la imagen siguiente se muestra el conjunto de prueba:



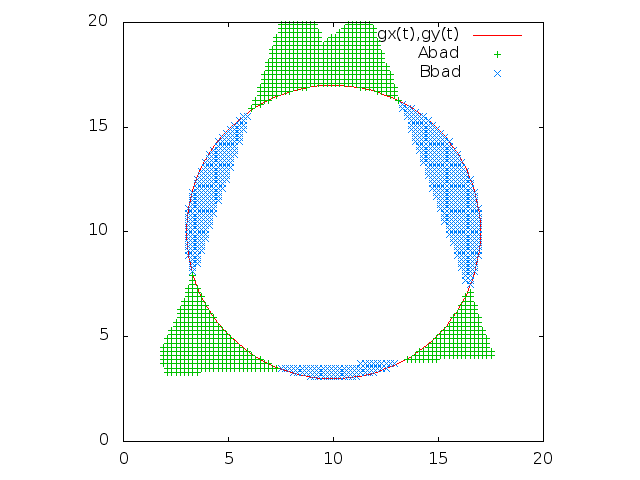
Los puntos azules representan aquellos puntos dentro del círculo y los morados aquellos que están por fuera del mismo.

Ahora se mostrarán tres gráficas, con el mismo conjunto de entrenamiento, pero con distinto número de neuronas de la capa intermedia.

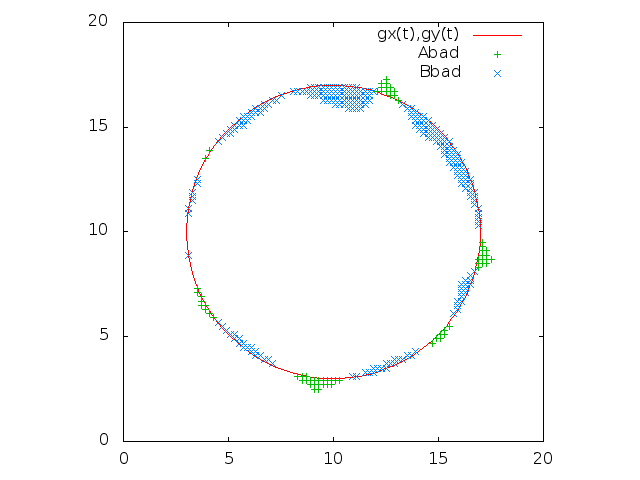
Con dos neuronas:



Con cuatro neuronas:



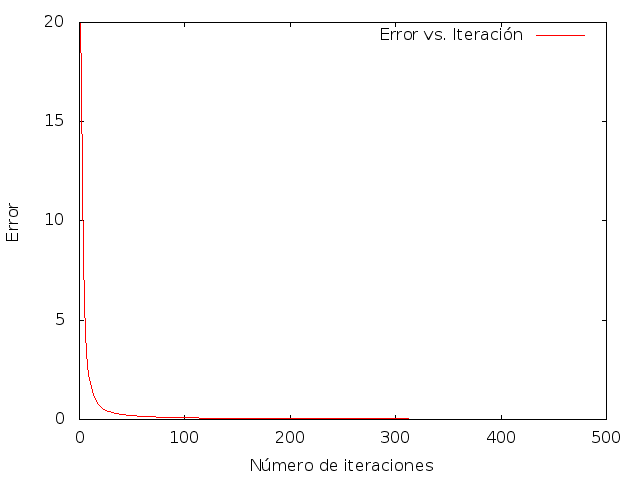
Con diez neuronas:

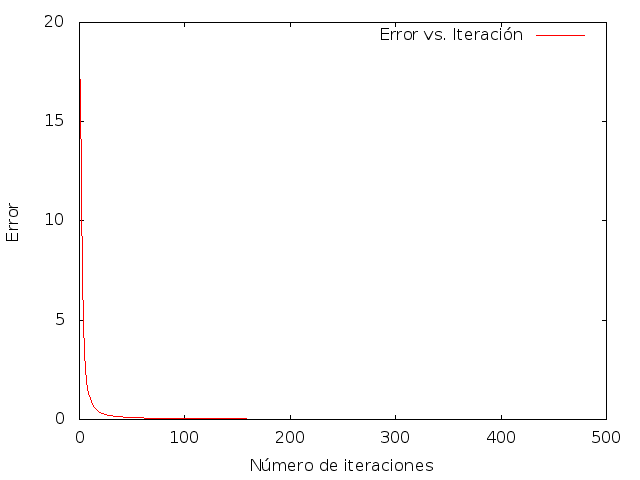


En cada una de las gráficas anteriores solo se gráfico los puntos que están mal clasificados. Los puntos verdes significan puntos que están fuera del círculo que se clasificaron como si estuviesen dentro del mismo, y los azules son aquellos que están dentro y se clasificaron como puntos que están por fuera del círculo.

Como se observa en las gráficas, existe una tendencia de que a mayor cantidad neuronas en la capa intermedia mejor es el resultado de las pruebas después del entrenamiento. Y esta observación justamente se evidencia para cada uno de los diferentes entrenamientos que se hizo con los distintos conjuntos de entrenamiento.

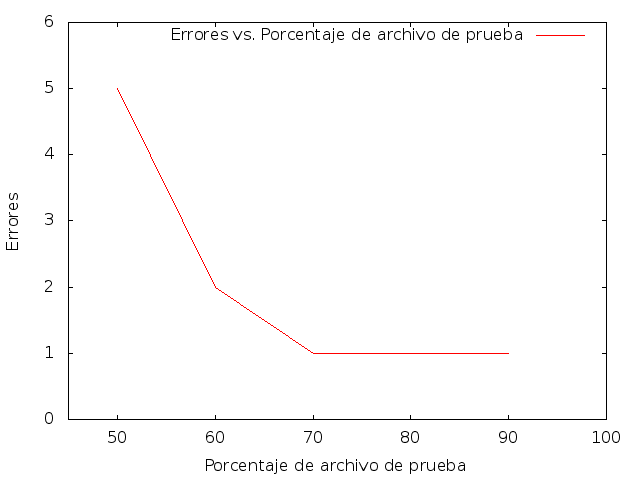
* + En el problema de la clasificación de las iris de forma binaria (es setosa o no), se quería graficar el porcentaje de error contra el porcentaje de fallas en la prueba, pero dado que en este problema el porcentaje de fallas siempre dio nulo, no parece relevante mostrar algún tipo de gráfica. Se concluyó que las iris setosa son separables de los otros dos tipos de iris. Sin embargo, para resaltar la importancia de la cantidad de neuronas en la capa intermedia de la red, y afianzar por qué su mejoría a una mayor cantidad de neuronas se tienen las siguientes gráficas:





En la primera se utilizaron cuatro neuronas en la capa intermedia y en la segunda diez para el mismo experimento. Como se observa, a mayor cantidad de neuronas la tendencia es que el error al entrenar converge más rápido a un error pequeño y este mismo comportamiento se evidenció en cada prueba hecha a la red en cada uno de los problemas.

* + En el problema de la clasificación de las iris en tres clases, la cantidad de falsas clasificaciones es similar para cualquier cantidad de neuronas y con algunas diferencias en el porcentaje de entrenamiento, como se muestra a continuación:



En cada experimento se tuvo como resultado que una cantidad mayor de entrenamiento tiende a dar un menor cantidad de clasificaciones falsas a la hora de realizar las pruebas.

La observación más importante, es que sin importar el número neuronas o el porcentaje de entrenamiento que se utilizó, siempre existía algún error de clasificación específicamente con las iris virginicas y las versicolor, con lo que se concluyó que no son separables.