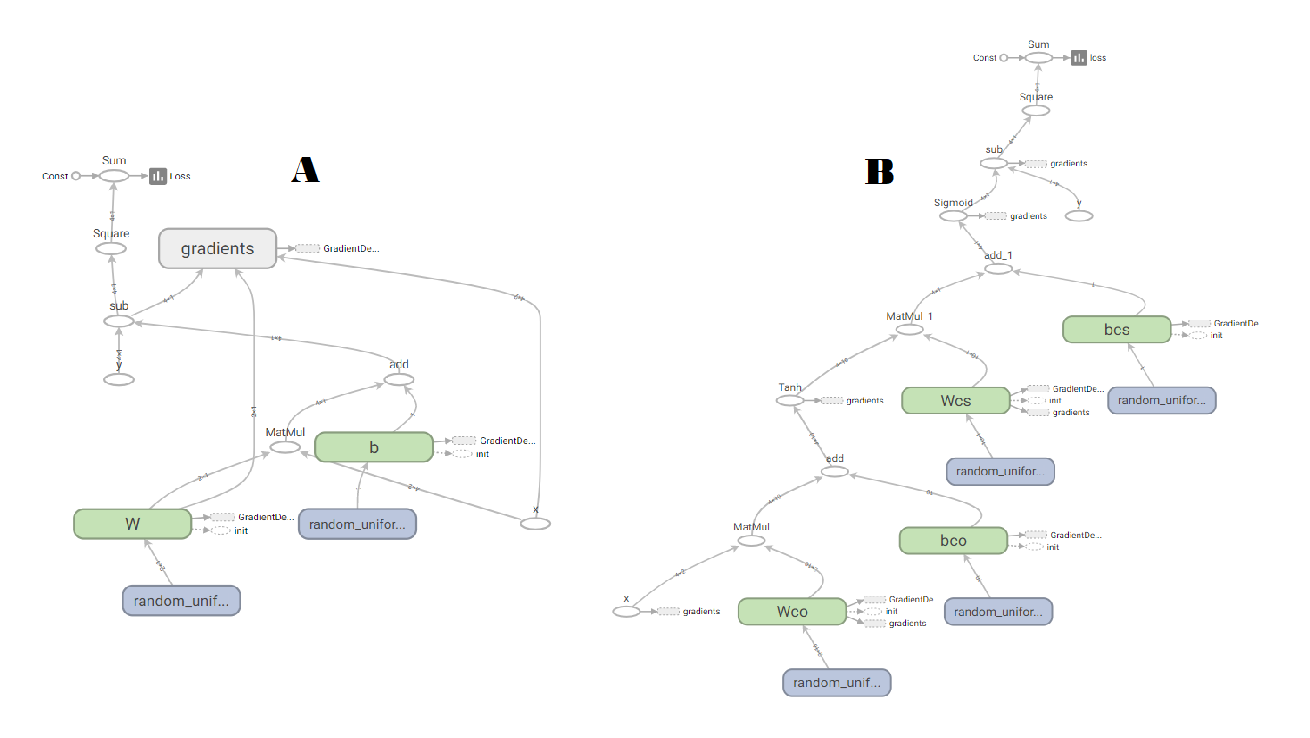
**1.**

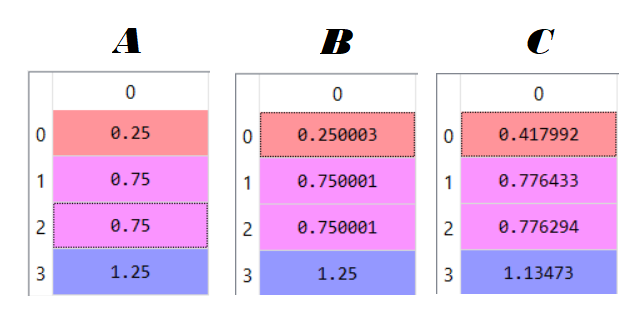
Para iniciar, se realizan 3 pruebas con una red adaline monocapa y otras 3 pruebas con una red adaline multicapa para simular la función lógica OR. En la *Ilustración 1* se puede ver los grafos obtenidos para cada red.



*ilustración 1: Grafos de la red Adaline monocapa (A) y la red Adaline multicapa (B)*

En la *ilustración 1* el grafo A representa una red Adaline monocapa de 2 entradas y una salida. El grafo B representa una red Adaline multicapa con 2 entradas, 10 neuronas en la capa oculta y una neurona de salida. En ambos grafos se puede apreciar las operaciones realizadas por cada red y el flujo que este sigue.

Posterior a esto, se realizaron 3 pruebas con la red Adaline monocapa, donde la *ilustración 2* corresponde a los valores de salida obtenido por dicha red variando el alfa o la constante de aprendizaje y *la ilustración 3* a la función de perdida obtenida por cada variación del alfa respectivamente.

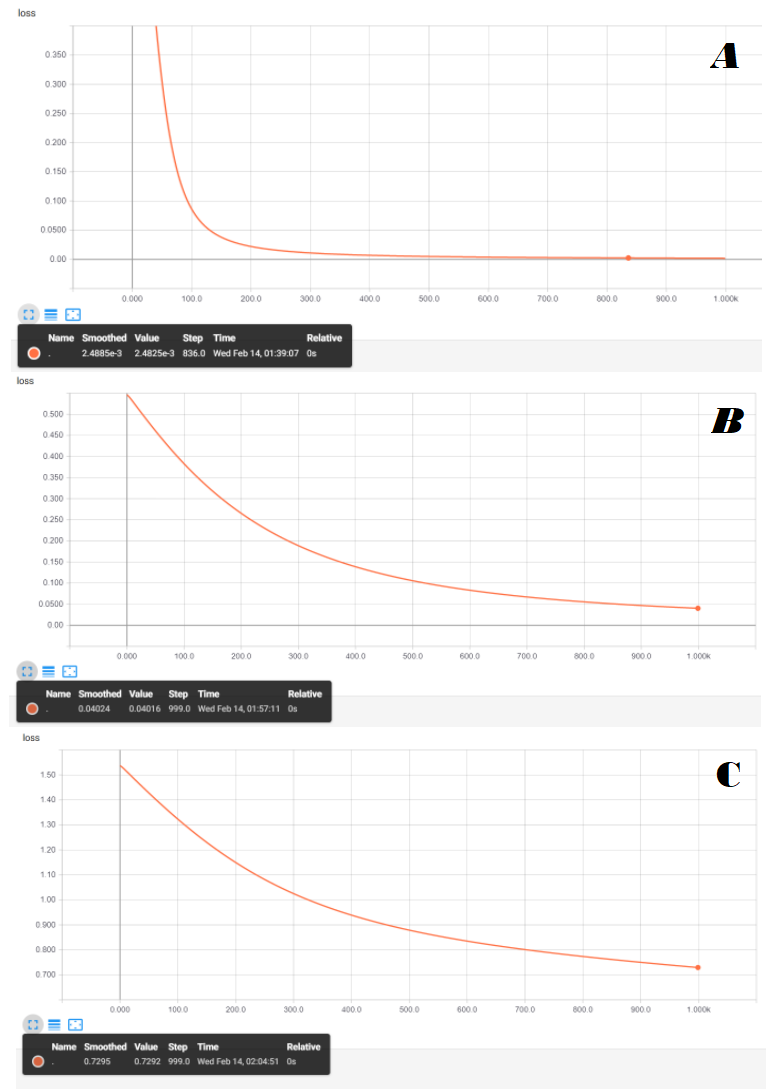


*ilustración 2: Datos de salida generados en la red Adaline monocapa para las respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

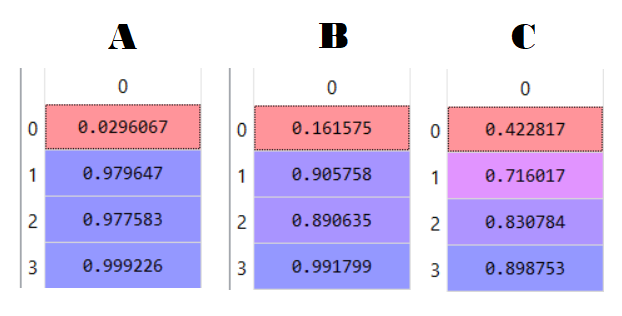


*ilustración 3: Función de pérdida generada para una función Adaline monocapa con sus respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

En la ilustración 3 se puede apreciar claramente que los mejores resultados obtenidos fueron con un alfa igual a 0.1 y 0.01, ya que al variarlo a 0.001, el error tienda a 0.2582, obteniendo un mayor chance de error. Asimismo, la misma prueba realizada en la red adaline monocapa se realizó en una red adaline multicapa como lo muestra la *ilustración 4 e ilustración 5* que muestran los valores de salida asociada a cada prueba con su respectiva función de perdida.



*ilustración 4: Función de pérdida generada para una función Adaline multicapa con sus respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*



*ilustración 5: Datos de salida generados en la red Adaline multicapa para las respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

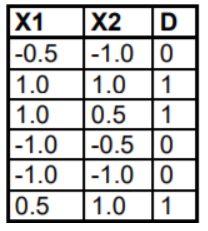
Dichas pruebas arrojan mejores resultados que la red adaline monocapa, puesto que la primera prueba con un alfa igual a 0.1, da un error de apenas , brindando a la salida la función lógica OR con valores muy aproximados a los esperados. Un punto a tener en cuenta es que a medida que se disminuyó el alfa, la red empezó a generar un mayor error respectivamente, esto debido a que las variaciones de los pesos eran cada vez más pequeñas debido a la disminución del alfa. Para corregir este problema, se podría haber modificado el número de épocas de 1000 a 10000, generando mejores resultados con un alfa mucho menor.

Lo anterior nos lleva a concluir que ambas redes pueden cumplir con dicha tarea, debido a que la función lógica OR es separable de manera lineal. Aún así, los mejores resultados son obtenidos por la red multicapa esto debido a su mayor capacidad de procesamiento, brindando un error de casi cero, a comparación de la red adaline monocapa, que el mejor resultado obtenido fue de 0.25.

**2.**

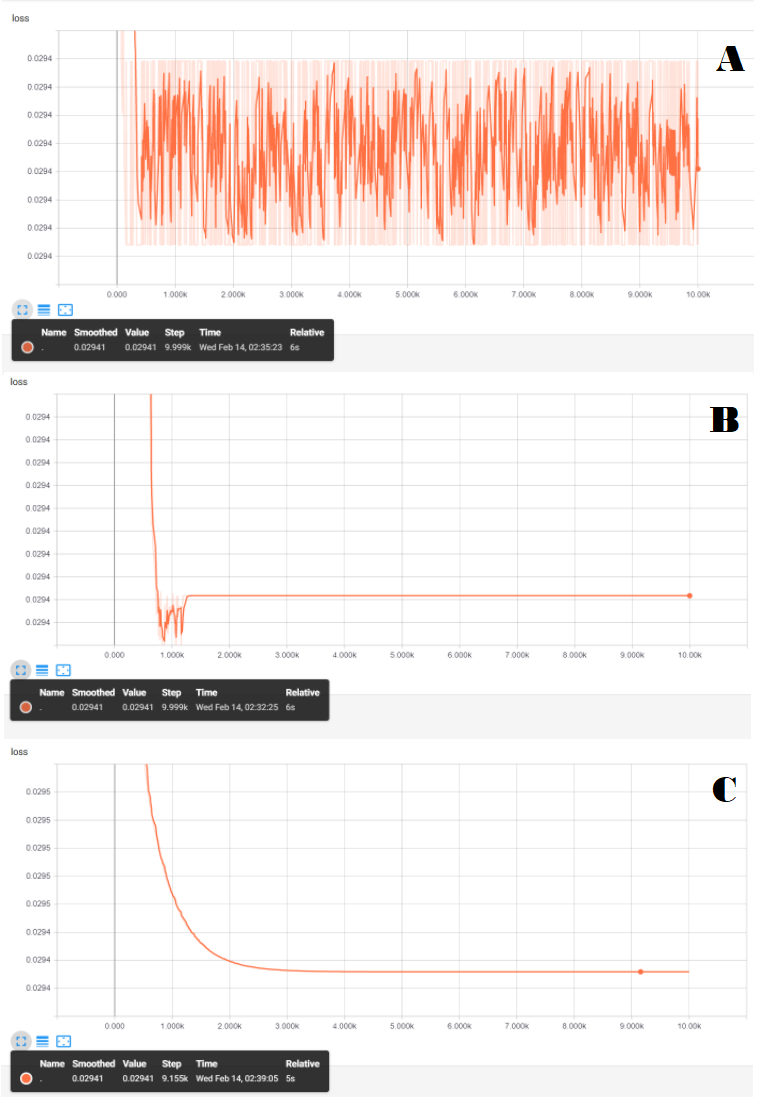
Para la función propuesta en la *ilustración 5* se realizó el mismo proceso realizado en el punto número uno, con la diferencia que se aumentó el número de épocas a 10000 y se usan las mismas variaciones de alfa (0.1, 0.01 y 0.001) utilizadas en el punto anterior. Cabe resaltar, que la función propuesta también puede ser separada o clasificada linealmente, lo cual brinda la oportunidad a la red adaline monocapa y multicapa de resolver el problema.

Para la red Multicapa, se utilizan 10 nodos en la capa oculta y una función de activación tanh en la capa oculta y Sigmoid para la capa de salida. Asimismo, no se repite la gráfica de los grafos puesto que esto son los mismo mostrados en la *ilustración 1*.

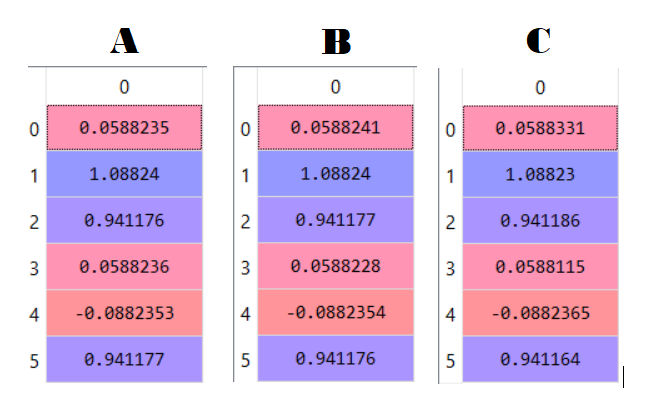


*ilustración 5: Función lógica a aprender con la Red*

En primera instancia, se realiza la simulación con la red adaline monocapa como se puede apreciar en la *ilustración 6 y 7*. Para la prueba A con el alfa 0.1, la función de perdida se ve algo ruidosa, pero no hay ningún problema con esto debido a que oscila entre el mismo valor (0.0294). En estas pruebas, se puede apreciar que al aumentar el número de épocas el resultado fue prácticamente el mismo en todas las pruebas, teniendo todo el mismo valor de error en la función de perdida y variaciones mínimas en el vector de salida dicha red.

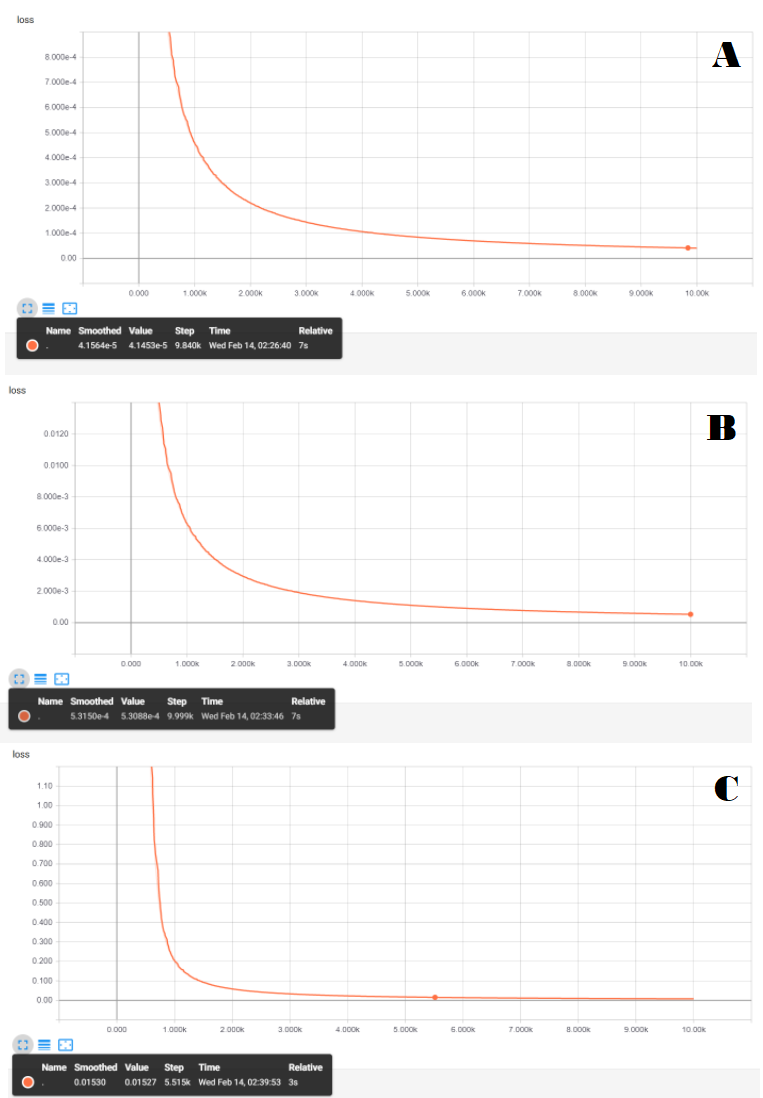


*ilustración 6: Función de pérdida generada para una función Adaline monocapa con sus respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

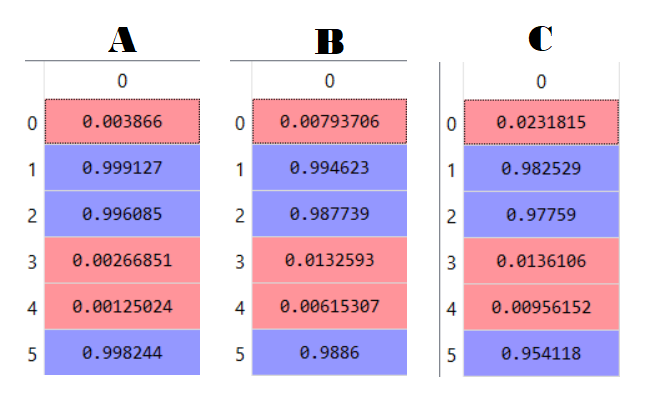


*ilustración 7: Datos de salida generados en la red Adaline monocapa para las respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

Por otra parte, a la hora de simular la red adaline multicapa, se resolvió también el problema, pero se obtuvieron mejores resultados en términos del error ya que la mejor prueba arrojó un error de con un alfa de 0.01, brindando una clasificación casi perfecta en este caso. Los resultados de las funciones de perdida y los vectores de salida de la red se pueden apreciar en la *ilustración 8 y 9* respectivamente.



*ilustración 8: Función de pérdida generada para una función Adaline multicapa con sus respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*



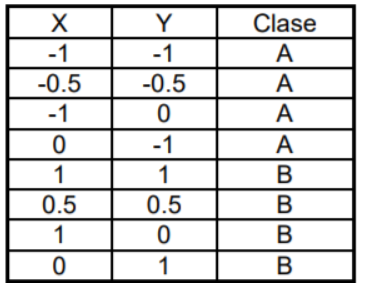
*ilustración 9: Datos de salida generados en la red Adaline multicapa para las respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

A modo de conclusión de este ejercicio, se concluye que ambas redes adaline cumplen correctamente con la tarea de clasificación, pero finalmente la red multicapa cumple mejor su tarea con un error de casi cero y resultados mucho más precisos como se puede comparar en la *ilustración 7 y 9* respectivamente. Por otra parte, se aprecia que al aumentar el número de épocas se reduce más la función de perdida, brindando la posibilidad de obtener mejores resultados a medida que se reduce el alfa.

**3.**

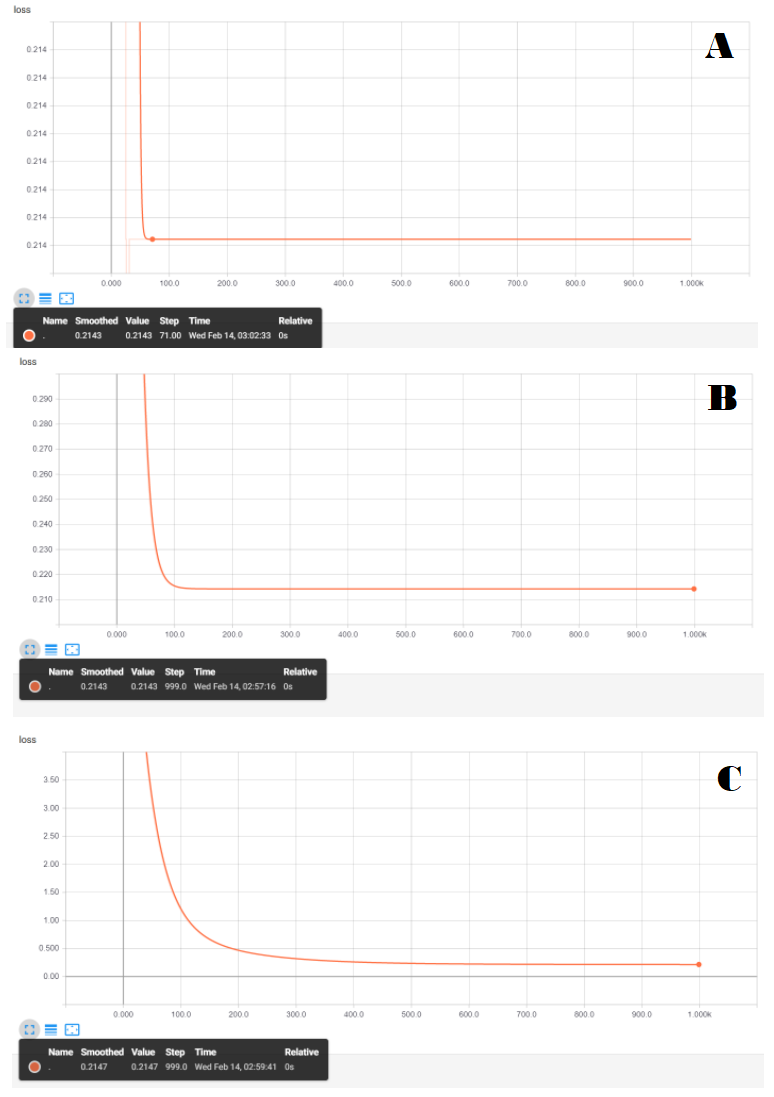
Dicha actividad propone clasificar la función propuesta en la *ilustración 10* donde A corresponde a 0 y B a 1, en esta ocasión se configura la fase de entrenamiento a 1000 épocas y se varía el número de neuronas en la capa oculta de la red multicapa a 8, asimismo, se utilizan las mismas funciones de activación para la red multicapa nombradas en el punto número 2.

Cabe resaltar que la función propuesta en la *ilustración 10*, es separable linealmente, por lo cual puede ser resuelta con la red Adaline multicapa y monocapa.

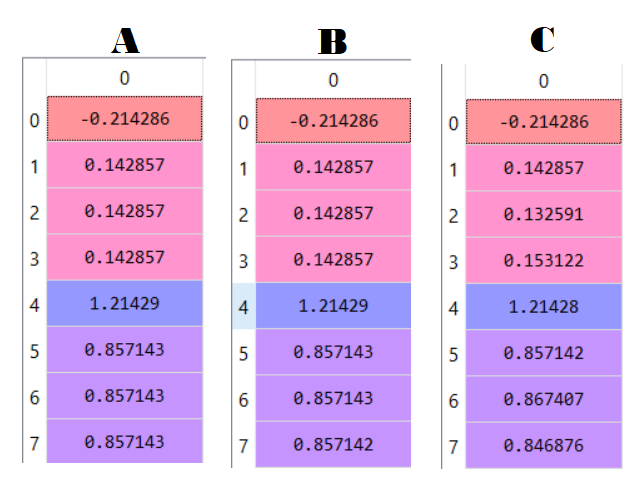
****

*ilustración 10: Función lógica a aprender con la Red*

Para la simulación de la red adaline monocapa, se simula bajo las condiciones indicadas al inicio de este inciso. La *ilustración 11 y 12* muestran los resultados obtenidos para dicha red, demostrando que no hubo muchas variaciones en la función de perdida ni en el vector de salida en el momento de variar el alfa para dicha red, brindando como mejor resultado un error de 0.2143 con un alfa igual a 0.1 y 0.01

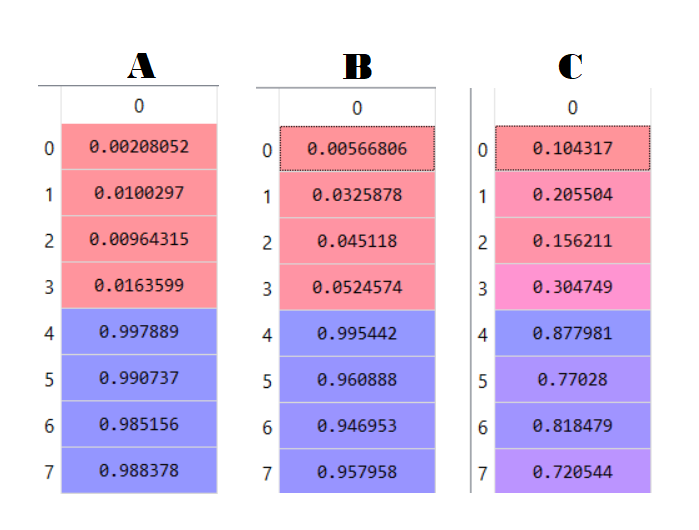
****

*ilustración 11: Función de pérdida generada para una función Adaline monocapa con sus respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

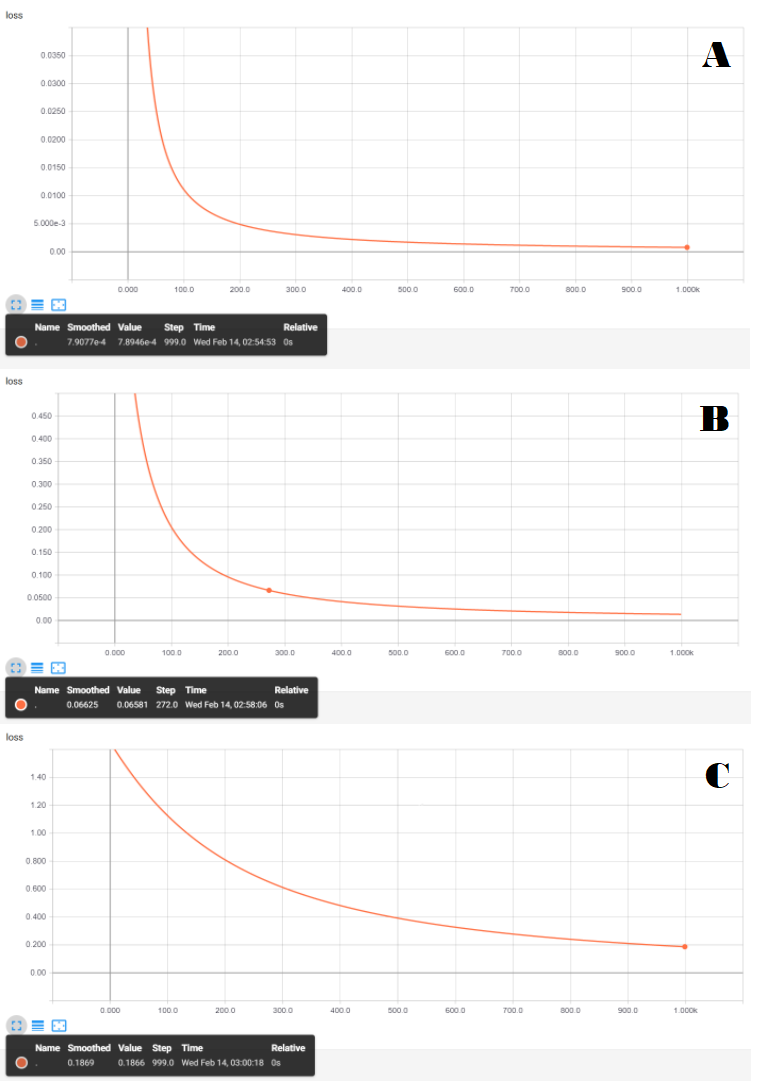
****

*ilustración 12: Datos de salida generados en la red Adaline monocapa para las respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

Igualmente se realizó la simulación con la red Adaline multicapa como se aprecia en la *ilustración 13 y 14*, brindando una vez mas mejores resultados que la red adaline monocapa, con un error de para un alfa de 0.1, en esta ocasión, a medida que se aumentaba el alfa, iba aumentando el error. Esto debido a que, al disminuir el número de épocas, las variaciones en los pesos eran muy pequeñas haciendo que la red no se entrenara de manera adecuada.



*ilustración 13: Datos de salida generados en la red Adaline multicapa para las respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

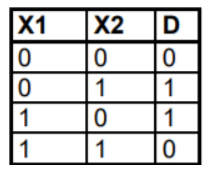


*ilustración 14: Función de pérdida generada para una función Adaline multicapa con sus respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.01 y C Alpha=0.001*

A modo de conclusión para esta prueba, se verifica una vez más que la red adaline multicapa es mejor para labores de clasificación brindando mejores resultados que la red adaline monocapa.

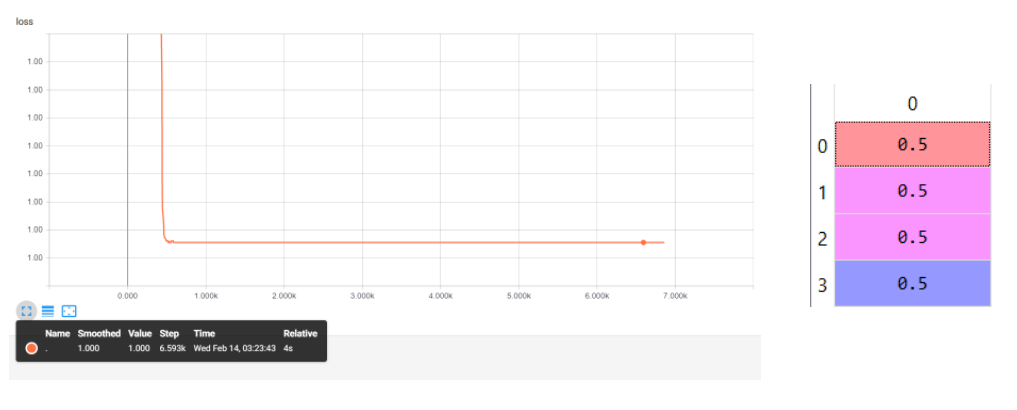
**4.**

En este ejercicio se propone realizar el entrenamiento de una red con la función lógica XOR de la *ilustración 15,* la cual no es separable linealmente y se demostrará realizando una simulación con la red adaline monocapa. La simulación se realizó con un total de 10000 épocas en cada red, asimismo, se vuelve a ajustar la red multicapa con 10 nodos en su capa oculta y se deciden variar las alfas en la simulación de la red multicapa con valores de 0.1, 0.5 y 0.8 respectivamente.



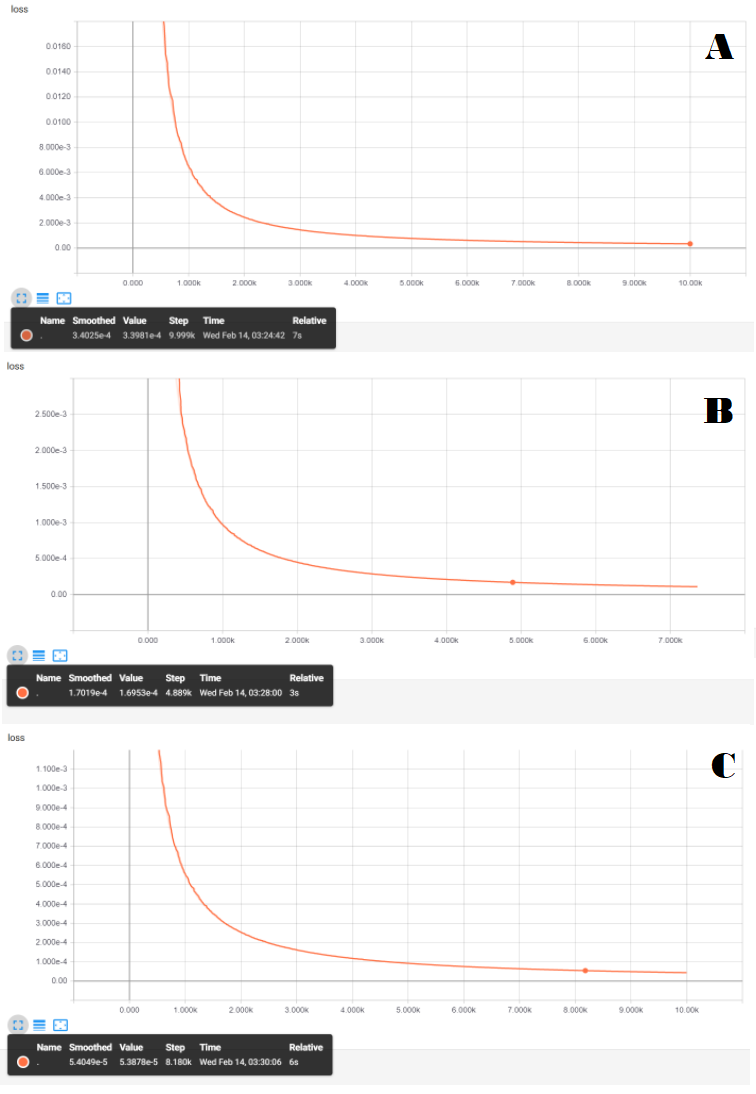
*ilustración 15: Función lógica XOR a aprender con la Red*

A la hora de simular la función lógica XOR con la red monocapa, se puede observar en la *ilustración 16* que los valores de salida son todos iguales y la función de perdida da un error de 1, generando el mayor error posible e indicando a su vez que la red no es capaz de generar la clasificación de dicha función, esto debido a como se indico anteriormente la función XOR no es separable linealmente.

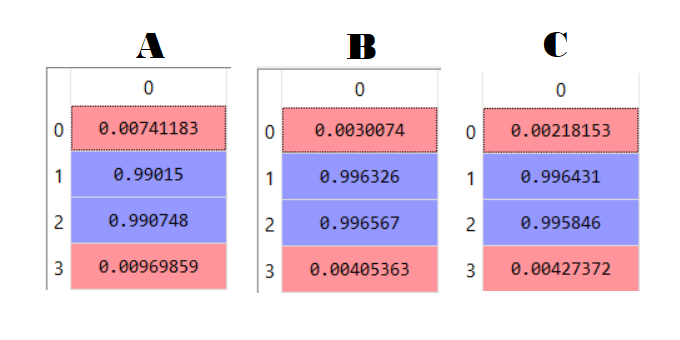


*ilustración 16: Función de pérdida generada para una función Adaline multicapa y valor de salida de la red generada con una variación del factor de aprendizaje: Alpha=0.1*

Por otra parte, cuando se simula la red adaline multicapa se puede observar en la *ilustración 17 y 18* que, si realiza el entrenamiento y clasificación de manera adecuada, con un error de con un alfa de 0.8. Brindando finalmente en el vector de salida los valores de la función lógica XOR con errores muy bajos en la función de perdida.



*ilustración 17: Función de pérdida generada para una función Adaline multicapa con sus respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.5 y C Alpha=0.8*

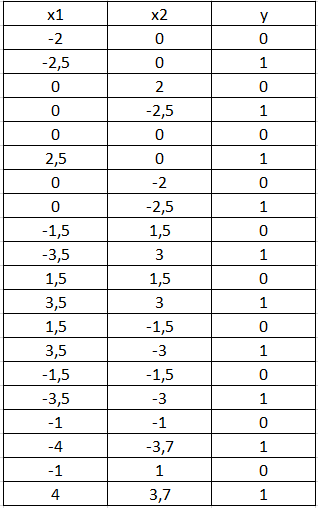


*ilustración 18: Datos de salida generados en la red Adaline multicapa para las respectivas variaciones del factor de aprendizaje: A Alpha=0.1, B Alpha=0.5 y C Alpha=0.8*

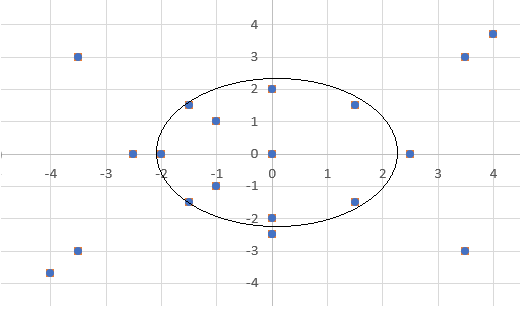
Cabe destacar, que en estas redes se decidió variar el alfa de manera ascendente y se obtuvieron mejores resultados con un total de 10000 épocas, no obstante, se debe tener cuidado ya que, si las variaciones de los pesos son muy abruptas haciendo que alfa tienda a uno, se puede generar que la red no quede entrenada de manera adecuada debida a dichas abruptas variaciones.

**5.**

Para el ejercicio propuesto, se genera la tabla contenida en la *ilustración 19* la cual genera dos regiones circulares, donde la región interna de círculos o la más cercana al origen corresponde a ceros y la región externa a esta a 1. Al ser un problema de separación no lineal, se resuelve el problema con una red Adaline superficial, donde se realizan 3 pruebas, una con un alfa igual a 0.5, 10 nodos en la capa oculta y 10000 épocas de entrenamiento, otra con un alfa igual a 0.1, 15 nodos en la capa oculta y 10000 épocas de entrenamiento y por último una prueba con un alfa igual a 0.8, 5 nodos en la capa oculta y 10000 épocas de entrenamiento. En la *ilustración 20* se puede apreciar la distribución de los puntos generados en el plano, donde los que están contenidos dentro de la circunferencia corresponden a los ceros, y los que están fuera a 1.

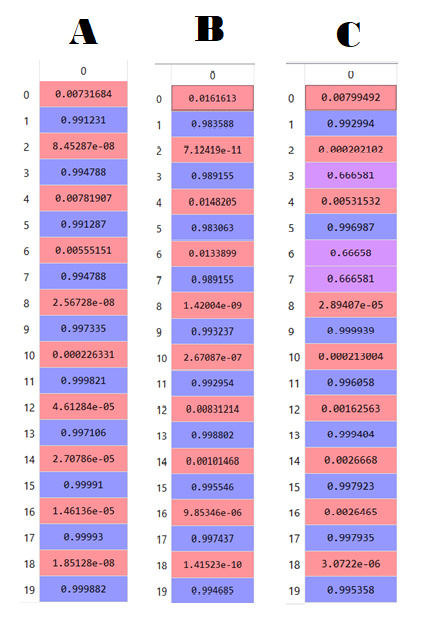


*ilustración 19: Función lógica a aprender con la Red*

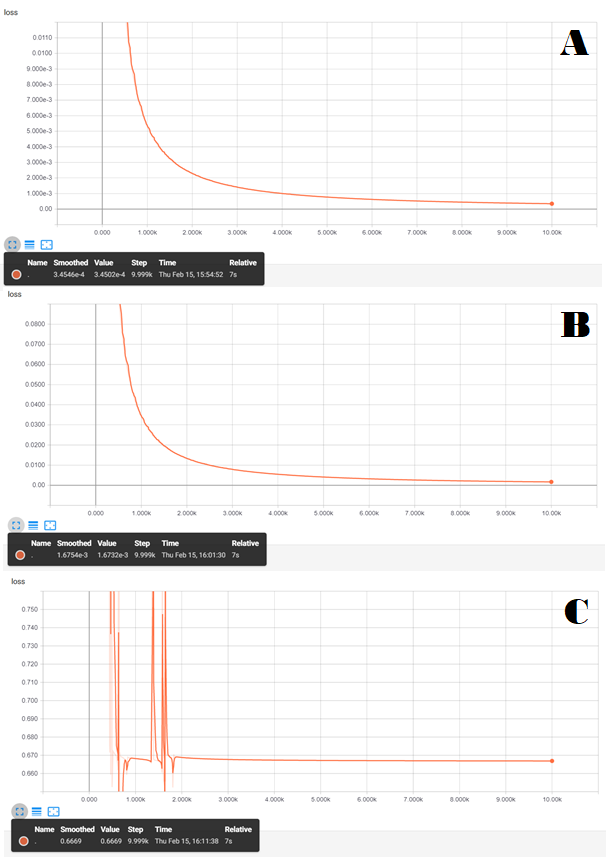


*ilustración 20: Puntos generados por la función lógica a aprender por la red*

Los resultados obtenidos en la simulación se enseñan en la *ilustración 21 y 22* donde se muestran los resultados obtenidos a la salida de cada prueba y las funciones de perdida obtenidas.



*ilustración 20: Datos de salida generados en la red Adaline multicapa con: A Alpha=0.5, 10 nodos y 10000 épocas; B A Alpha=0.1, 15 nodos y 10000 épocas; y C A Alpha=0.8, 5 nodos y 10000 épocas*



*ilustración 21: Función de pérdida generada para una función Adaline multicapa con: A Alpha=0.5, 10 nodos y 10000 épocas; B A Alpha=0.1, 15 nodos y 10000 épocas; y C A Alpha=0.8, 5 nodos y 10000 épocas*

Se puede analizar que los mejores resultados obtenidos fueron para un alfa de 0.1 y 15 nodos en la capa oculta, dando un error de lo cual es casi cero, llevándonos a concluir que la red adaline superficial es capaz de realizar de manera adecuada la separación de las 2 circunferencias con 15 nodos en la capa oculta. Por otra parte, se pudo observar que a medida que se reducía el número de nodos y el alfa, el error tendría a ser muy grande como se enseña en la prueba C donde el error es de 0.67, brindando una mala clasificación de los datos, lo anterior debido a que al reducir los nodos en la capa oculta da menos procesamiento de información en dicha capa y es más duro realizar una correcta clasificación de los datos, así mismo, si se generan muchos nodos en la capa oculta, se aumenta considerablemente la capacidad de cómputo y puede que la red quede sobreentrenada. Es por esto que el mejor resultado se obtiene con 15 nodos, un alfa igual a 0.1 y un alto número de épocas para el entrenamiento.