



IPN-UPIITA

Redes Neuronales

Reporte R02

Dr. Rafael Martínez Martínez

Academia de sistemas

ramartinezr@ipn.mx

Instrucciones:

- Cada problema/ejercicio debe tener procedimiento ordenado y completo que justifique adecuadamente la respuesta anotada.
 - Si falta el procedimiento o este no justifica la respuesta anotada entonces el problema vale 0 puntos aunque la respuesta sea correcta.
-

Contenido

Problema 1 (50 puntos)	1
Problema 2 (50 puntos)	2

Problema 1 (50 puntos)

Walter Pitts salio de las calles al MIT, pero no pudo escapar de sí mismo.

La vida de Walter Pitts pasó de un vagabundo fugitivo, a pionero de la neurociencia del MIT, a alcohólico retraído.

En 1943 Warren McCulloch y Walter Pitts introdujeron una de las primeras neuronas artificiales. La referencia de dicho trabajo es la siguiente:

W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," Bulletin of Mathematical Biophysics, Vol. 5, pp. 115–133, 1943.

Este artículo introduce el primer modelo matemático de una neurona, en la cual la suma ponderada de las señales de entrada es comparada con un umbral que determina la salida de la neurona. Cuando la suma es más grande o igual al umbral, la salida es 1. Cuando la suma es menor al umbral, la salida es 0. A diferencia de las redes biológicas, los parámetros de su red (pesos) se tenían que diseñar y no proporcionaron un método de entrenamiento. Sin embargo, fue el precedente sobre un campo interesante entre la conexión de la biología y las computadoras digitales.

A finales de la década de los 50's del siglo pasado, Frank Rosenblatt y otros investigadores desarrollaron una clase de redes neuronales llamada perceptrones:

F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," Psycho- logical Review, Vol. 65, pp. 386-408, 1958.

Su modelo era parecido al de McCulloch y Pitts, pero además proponían una regla de aprendizaje para resolver el problema de reconocimiento de patrones. Posteriormente se mostró que este aprendizaje estaba limitado a resolver problemas de reconocimiento linealmente separables. Y no fue hasta la década de los 80's del siglo pasado, que esta limitación se resolvió utilizando múltiples capas de perceptrones.

La colaboración de Warren McCulloch y Walter Pitts va más allá de la profesional. En el artículo de divulgación [The Man Who Tried to Redeem the World with Logic](#), se resume un poco de la desgarradora vida de Pitts. En la lectura se podrá observar como Pitts colaboró con algunas de las grandes mentes del siglo XX. Para este problema es necesario realizar la lectura del artículo de divulgación mencionado. Para el control de dicha lectura, se responden las siguientes preguntas (lo principal es tu reflexión, que no se pide, sobre el contenido de la lectura):

1. ¿Qué edad tenía Pitts cuando le escribió a Bertrand Russell sobre los errores en su libro?
2. ¿Quién presentó a Pitts con McCulloch?
3. ¿Quiénes conformaban el núcleo del grupo conocido como *los cibernéticos*?
4. ¿Cuál fue el motivo por el cual Norbert Wiener dejó de tener comunicación con Pitts?
5. ¿Cuáles fueron los eventos que llevaron a Pitts a hundirse en depresión, y quemar su tesis doctoral y sus escritos?
6. ¿Consideras que si Pitts hubiera tenido una situación diferente a la que tuvo en el lugar donde nació, podría haber superado los eventos que lo llevaron a la depresión? Explica

Problema 2 (50 puntos)

La siguiente [aplicación web](#), permite elegir gráficamente una topología de red (feedforward) para resolver problemas de clasificación y regresión preestablecidos.

La Figura 1 muestra la elección de una topología para resolver un problema

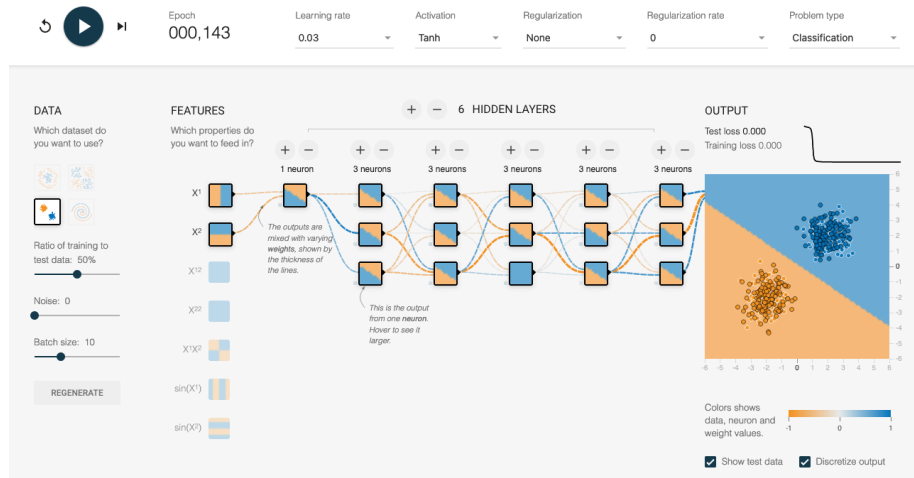


Figura 1. Red elegida par resolver el problema de clasificación indicado

(porporcionado en la aplicación) de clasificación. Se han elegido algunas capas ocultas¹ y los demás parámetros de la red. Al comenzar en entrenamiento y detenerlo después de 143 epocas, los valores de los pesos se pueden consultar colocando el puntero sobre las líneas de cada capa.

Si bien se tienen dos entradas en cada problema de clasificación (en esta aplicación), la aplicación nos permite generar 5 entradas adicionales construidas con los dos datos de entrada, además tiene la opción de tomar estas entradas *artificiales* como información adicional para entrenar la red. Es decir, se puede pensar que dado un problema con m entradas se puede transformar a un problema de $n + m$ entradas, donde n es el número de datos construidos *artificialmente* transformando los m datos iniciales. En el caso de la aplicación $n = 5$.

Conteste las siguientes preguntas

1. ¿A qué se refiere que una red sea feedforward?
2. ¿Cuántos problemas de clasificación y cuántos de regresión tiene preestablecidos la aplicación?
3. Investiga a qué se refieren los siguientes elementos que aparecen en la aplicación²:
 - a. Learning rate
 - b. Regularization
 - c. Regularization rate

¹Por el momento pensemos que si se tiene más de una capa, el entrenamiento se lleva a cabo de alguna manera, que por el momento no nos importa especificar.

²Explicación breve, si lo consideras necesario anota ecuaciones como ejemplo

- d. Ratio of training to test data
 - e. Batch size
 - f. Test loss
 - g. Training loss
4. Es claro que la topología elegida en la Figura 1 *está sobrada*. Pues tenemos un problema de clasificación linealmente separable. Resuelve este problema con una capa, una neurona, las dos estradas *naturales*, y la elección de tu preferencia de los demás parámetros (función de activación, Learning rate, etc.), no es necesario justificar la elección de los parámetros.
- a. Solo reporta la imagen correspondiente, donde se pueda apreciar que se llevó acabo el entrenamiento de forma exitosa.
 - b. ¿Cuántas épocas de entrenamiento fueron necesarias (para esto utiliza el botón a un costado de *play*)?
5. **Juega con la aplicación.** Resuelve los problemas de clasificación restantes, utilizando la menor cantidad de capas posibles y la menor cantidad de neuronas posibles en cada capa³. Todos los demás parámetros se pueden elegir a tu consideración (entradas, función de activación, Learning rate, etc.) no es necesario justificar la elección de los parámetros. Reporta solo las imágenes correspondientes, donde se pueda apreciar que se llevó acabo el entrenamiento de forma exitosa.

³Varias topologías resolverán el problema, queremos la más sencilla que puedas construir, en este caso nos referimos a sencilla en el sentido de minimizar el número de capas y neuronas