

Redes neuronales para la caracterización de yacimientos: del mito a las aplicaciones prácticas.

Rafael E. Banchs, UCAB, Caracas, Venezuela

En los últimos años, a medida que han proliferado las aplicaciones tecnológicas de la llamada inteligencia artificial, algunos mitos existentes en torno a estas tecnologías emergentes han ido desapareciendo. Sin embargo a estas alturas, la definición precisa de inteligencia sigue siendo incierta y la de inteligencia artificial más aún. Aunque su inicio se remonta a más de medio siglo cuando los psicólogos presentaron sus primeros modelos del cerebro y el aprendizaje (McCulloch y Pitts, 1943) no fue sino hasta finales de los 80, cuando la tecnología computacional estuvo lista, que la aplicación práctica de estos conceptos a problemas de la vida diaria, la industria y la ciencia comenzó a materializarse.

Primero fueron los sistemas expertos (Waterman, 1986) y luego llegaron las redes neuronales (Fahlman y Hinton, 1987), la fascinación del público no había terminado aún cuando llegaron los algoritmos genéticos (Holland, 1992) seguidos de la lógica difusa (McNeill y Freiberger, 1993). Hoy en día se habla de cualquier combinación y/o permutación de estos términos y es una forma segura y rápida de vender tecnología. El problema de todas estas tecnologías radica en que para la mayoría del público en general, e incluso para un gran número de sus usuarios, continúan siendo cajas negras que siempre arrojan una solución.

Este trabajo tiene un doble objeto; en primer lugar, pretende desmitificar un poco el concepto de las redes neuronales mediante la presentación de los aspectos más fundamentales de su teoría y la discusión de sus ventajas y desventajas como métodos de inferencia; en segundo lugar, presenta un resumen de las aplicaciones comerciales más populares que usan redes neuronales en problemas de caracterización de yacimientos y dos aplicaciones específicas desarrolladas y aplicadas en yacimientos venezolanos.

Nociones básicas sobre redes neuronales

Una red neuronal artificial es un modelo matemático entrenable que permite encontrar relaciones funcionales entre dos conjuntos de datos. En este artículo se pretende explicar con lujo de detalles el significado de esta definición.

El concepto de neurona artificial se deriva de los primeros modelos que los psicólogos presentaron sobre el funcionamiento de las neuronas biológicas en cuanto a dos aspectos básicos: la respuesta a estímulos y el aprendizaje. Un perceptrón, o neurona artificial, es la unidad de procesamiento fundamental de toda red neuronal. La red, como su nombre lo indica, constituye la interconexión de varias de estas unidades de procesamiento. La estructura de la interconexión, así como los distintos niveles estructurales en que se pueden disponer los perceptrones dentro de una red definen lo que se denomina la arquitectura de la

red. Por otra parte, el tipo de información provista en los datos y la naturaleza del problema en cuestión, determina la forma de aprendizaje que debe ser utilizada y a su vez la arquitectura más idónea. Independientemente de la arquitectura, el tipo de aprendizaje y los datos, el perceptrón como unidad básica de procesamiento es básicamente el mismo para cualquier arquitectura y/o tipo de red. La figura 1 presenta una neurona biológica y una neurona artificial indicando cada una de sus partes constitutivas.

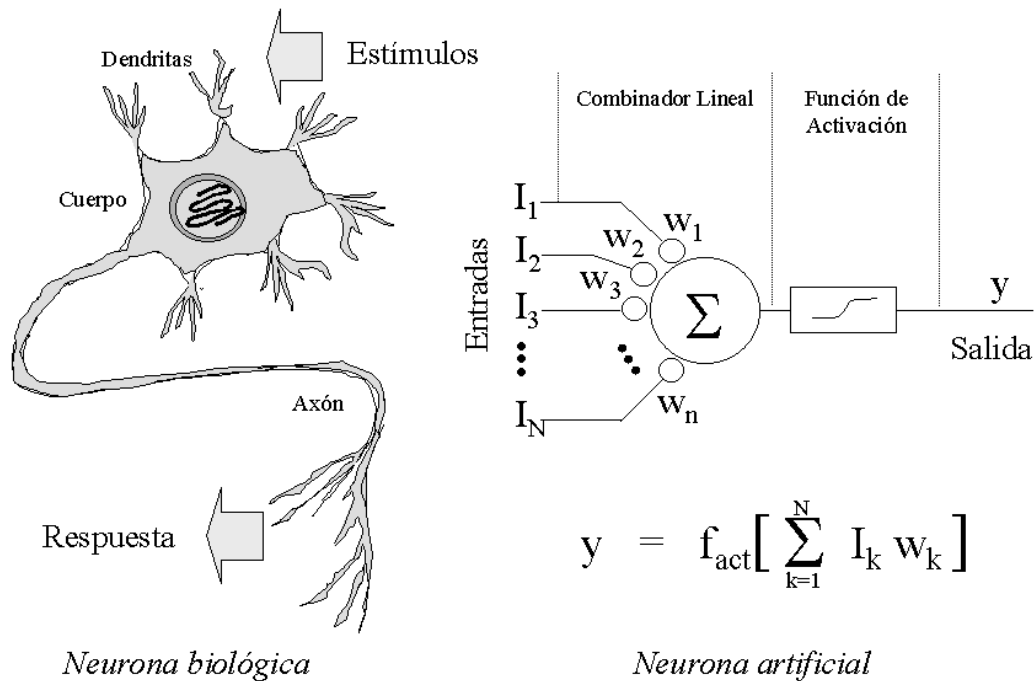


Figura 1. Neuronas artificial y biológica

Cómo se observa de la figura, tanto la neurona biológica como la artificial son sistemas que reciben varias entradas ó estímulos y producen una salida o respuesta. En una red neuronal, las salidas de unas neuronas constituyen entradas de otras neuronas. La interconexión de una de las terminaciones de un axón con una dendrita de otra neurona se denomina sinapsis.

En el funcionamiento de una neurona biológica, la preponderancia relativa de cada estímulo en la generación de la respuesta depende del número de neurotransmisores y neuro-receptores existentes en cada sinapsis y se ha demostrado que las sinapsis se refuerzan o debilitan según sea la historia de su actividad; ésta es, como veremos más adelante, la base del aprendizaje Hebbiano que debe su nombre al neuropsicólogo Hebb quién fue el primero en plasmar estas ideas en su obra 'The Organization of Behavior' (1949).

En el modelo artificial de una neurona, la fuerza de una sinapsis se representa mediante un peso que multiplica al valor de la entrada o estímulo. Finalmente, como se ilustra en la figura 1, la respuesta de una neurona artificial está definida

por el resultado de aplicar una función de activación (generalmente no lineal) a la suma pesada de todas las entradas de la neurona. De esta forma, el modelo de neurona artificial está compuesto por un combinador lineal seguido de una función de activación.

Redes neuronales y filtros

Consideremos por un momento el modelo más simple de neurona que podamos definir. Sería aquella en la que la función de activación está definida por la función identidad $y=x$. El modelo se reduce al combinador lineal para el cual la salida está dada por la suma pesada de sus entradas:

$$y = \sum_{k=1}^N I_k w_k$$

Ecuación 1

La ecuación 1 es exactamente la misma ecuación de un filtro lineal e invariante en tiempo de respuesta impulsiva finita (Oppenheim y Schaffer, 1989). Si conocemos la respuesta deseada (d) para un conjunto de datos de entrada (I_k), podemos definir el error como la diferencia entre la respuesta deseada y la respuesta del filtro:

$$e = d - y$$

Ecuación 2

Y podemos buscar un conjunto de pesos (w_k) tal que minimice el error. Una forma de hacer dicha búsqueda es mediante el uso de una función de costo definida a partir del error cuadrático medio, la cual se define como el valor esperado (sobre el universo de datos) del cuadrado del error, como se muestra a continuación:

$$ecm = \frac{1}{2} E [e^2]$$

Ecuación 3

De esta forma, la ecuación 3 es una función de los datos de entrada, sus respuestas deseadas y los pesos del filtro, y su valor mínimo ocurre cuando el valor esperado para el error es cero. Calculando su gradiente e igualándolo a cero, podemos calcular los valores de los pesos que minimizan el error. Este procedimiento no es nada más que el problema de filtraje lineal óptimo resuelto por Wiener a mediados del siglo pasado (Widrow y Stearns, 1985); y en su honor, a este conjunto de pesos óptimo se le denomina filtro Wiener. El cálculo de estos pesos óptimos no es siempre sencillo ya que una buena estimación del valor esperado del cuadrado del error no es siempre posible. Existen métodos recursivos, tales como el *Steepest Descent* y el *Least-Mean-Square*, que permiten

la búsqueda de los pesos óptimos de una forma más robusta (Haykin, 1996). En general, cualquier método de optimización, incluyendo algoritmos genéticos y *simulated annealing*, puede ser usado.

En el contexto de las redes neuronales, los conjuntos de datos de entrada para los cuales se conocen sus salidas deseadas se denominan los datos de entrenamiento; y entrenar una red neuronal no es otra cosa más que buscar un conjunto de pesos óptimos para representar de la mejor manera posible una relación funcional entre el espacio de los datos de entrada y el espacio de los datos de salida. El gran poder de las redes neuronales radica en la función de activación. La inclusión en el modelo de la neurona artificial de una función de activación no lineal permite la representación de relaciones funcionales no lineales entre el espacio de los datos de entrada y el espacio de los datos de salida. De esta forma una red neuronal no es más que un gran filtro no lineal construido a partir de la combinación de pequeños filtros no lineales.

Clasificación con redes neuronales

Veamos con un ejemplo sencillo, como una neurona artificial se puede utilizar en un problema de clasificación en un espacio de dos variables. Considere el ejemplo de la figura 2.

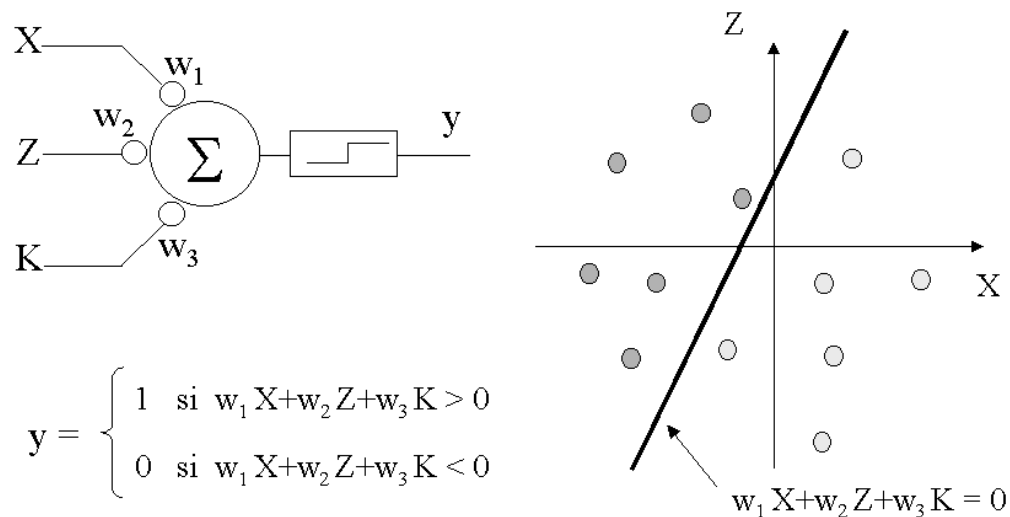


Figura 2: Problema de clasificación simple con una neurona.

Cómo puede observarse de la figura 2, se implementa un clasificador con una neurona de tres entradas y una función de activación del tipo escalón (función de heaviside). Este modelo de neurona se denomina el modelo de McCulloch-Pitts (Haykin, 1994). Cuando la entrada de la función de activación es negativa, la respuesta de la neurona será cero (neurona inactiva); y cuando la entrada de la función de activación es positiva, la respuesta de la neurona será uno (neurona activa). Recordemos que la entrada a la función de activación está dada por la

salida del combinador lineal, y en este caso en particular, como se ilustra en la figura, está dada por la suma pesada de las entradas X , Z y K . Nótese como la condición umbral ($w_1X+w_2Y+w_3K=0$) define una recta en el plano X, Z separándolo en dos semiplanos. El entrenamiento de este clasificador consiste en encontrar un conjunto de valores para w_1, w_2, w_3 y K , tales que las dos clases definidas por los puntos grises y los puntos blancos queden confinadas en semiplanos diferentes.

Para ilustrar el poder de combinar elementos no lineales en un problema de clasificación, consideremos el siguiente ejemplo.

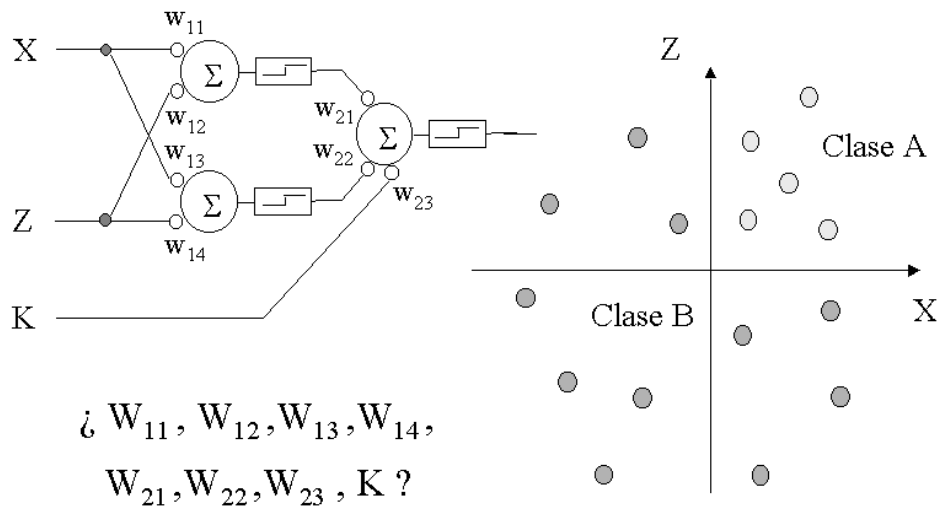


Figura 3: Clasificación con una red neuronal de dos capas

En el ejemplo ilustrado en la figura 3, se implementa un clasificador con una red neuronal de tres neuronas organizadas en dos niveles o capas. Esta configuración le da a la red la potencialidad de discriminar zonas complicadas en el plano X, Z . En este ejemplo particular se desea discriminar el primer cuadrante del resto del plano. La red presentada en el ejemplo es capaz de resolver este problema, el cual es imposible de ser resuelto con una sola neurona como en el caso del ejemplo de la figura 2. La solución del problema es relativamente sencilla, cada una de las dos neuronas de la primera capa (o capa de entrada) se encargan de dividir el plano X, Z en dos semiplanos, tal y cómo se ilustró en el ejemplo de la figura 2. Por otra parte, la neurona de la segunda capa (o capa de salida), se encarga de interceptar las soluciones generadas por las neuronas de la capa de entrada logrando así discriminar el primer cuadrante del plano X, Z .

El problema planteado en la figura 3 admite más de una solución válida. De hecho, existen infinitud de combinaciones de valores para $W_{11}, W_{12}, W_{13}, W_{14}, W_{21}, W_{22}, W_{23}$ y K que resuelven el problema de clasificar las clases A y B en el espacio X, Z ; se invita al lector a encontrar algunas de ellas. En general, la arquitectura de red presentada en la figura 3, permite discriminar cualquier área del plano definida por la intersección de dos rectas.

Estimación con redes neuronales

Al igual que el problema de clasificación, el problema de estimación puede ser abordado con una red neuronal. En este caso, de la misma forma en que una serie de funciones puede ser usada para aproximar otra función, la red neuronal puede aproximar una función mediante la combinación de aproximaciones individuales generadas por cada uno de los perceptrones de la red. Tal es el caso de las denominadas RBF: Radial Basis Functions (Haykin, 1994).

En la figura 4 se ilustra un problema sencillo de estimación mediante el uso de una red neuronal del tipo RBF. En dicho ejemplo, se pretende aproximar una función lineal ($5Y=7X+21$) en el intervalo $-2 < X < 2$. Esta aproximación se realiza con una red neuronal de dos capas, tal y como se ilustra en la figura. La primera capa esta constituida por seis neuronas que poseen dos entradas y una función de activación gaussiana:

$$f(s) = e^{-\frac{s^2}{k}}$$

Ecuación 4

donde s , corresponde a la suma pesada de las entradas X y Z ; y k es un factor de escala predefinido (para el problema en cuestión se uso $k=16/25$). La segunda capa está constituida por una única neurona con una función de activación lineal (función identidad en este caso).

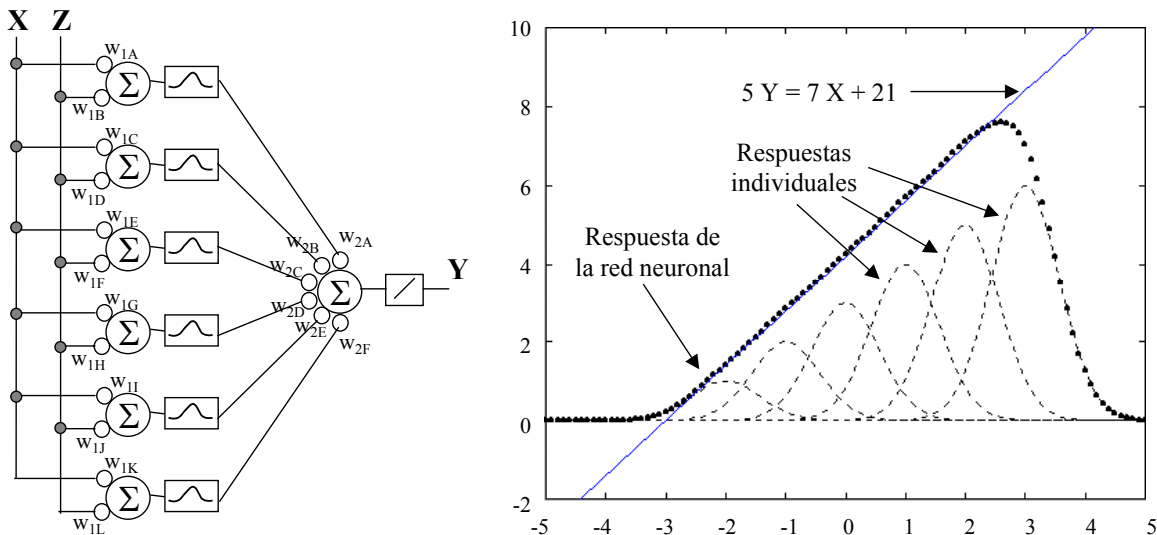


Figura 4: Estimación con una red neuronal del tipo RBF.

Como se observa en la figura, la respuesta de la red neuronal aproxima bastante bien a la función lineal en el intervalo $-2 < X < 2$. Dicha respuesta es la suma pesada de las respuestas individuales de cada una de las seis neuronas de la primera

capa. Cada una de estas neuronas contribuye al resultado total en una forma localizada según el valor de Z , X y sus pesos. En cada neurona, Z y su peso asociado definen el centroide correspondiente a la gaussiana.

En el ejemplo mostrado en la figura 4, el valor de Z fue 1 y los pesos fueron: $W_{1A}=W_{1C}=W_{1E}=W_{1G}=W_{1I}=W_{1K}=1$, $W_{1B}=-2$, $W_{1D}=-1$, $W_{1F}=0$, $W_{1H}=1$, $W_{1J}=2$, $W_{1L}=3$, $W_{2A}=1$, $W_{2B}=2$, $W_{2C}=3$, $W_{2D}=4$, $W_{2E}=5$ y $W_{2F}=6$.

De la misma manera en que se aproximó una función lineal en el ejemplo anterior, es posible aproximar cualquier función genérica en una, dos o más dimensiones usando arquitecturas de redes neuronales similares a la de la figura 4, y diversos tipos de funciones de activación.

Algoritmos y paradigmas de aprendizaje

En los ejemplos presentados anteriormente se ilustró cómo las redes neuronales pueden ser utilizadas para problemas de clasificación y estimación. Esto se ilustró con una arquitectura particular denominada perceptrón de capas múltiples y los pesos de las neuronas fueron propuestos o calculados en forma intuitiva. Este procedimiento, en cualquier aplicación práctica, carece de validez y sentido ya que la determinación directa de los pesos implica el conocimiento de la solución del problema. La verdadera utilidad práctica de una red neuronal como método de inferencia, radica en el hecho de que para un problema dado es posible determinar los pesos de las neuronas de una red mediante el uso de un procedimiento específico denominado algoritmo de aprendizaje (Haykin, 1994).

El proceso de aprendizaje de una red neuronal se lleva a cabo mediante la ejecución de ciertos cambios en la configuración de la red, específicamente, sus pesos. Desde el punto de vista algorítmico, estos cambios se realizan a modo de una actualización que se representa de la siguiente forma:

$$W_{ij}(n+1) = W_{ij}(n) + \Delta W_{ij}(n) \quad \text{Ecuación 5}$$

donde n constituye in índice temporal, y ΔW_{ij} representa el valor de la actualización en un momento determinado. Una regla o conjunto de reglas que permite calcular el valor de la actualización de los pesos es lo que se denomina un algoritmo o regla de aprendizaje. Aunque hay muchos algoritmos de aprendizaje diferentes, existen tres reglas básicas que son las más comúnmente usadas: el aprendizaje Hebbiano, el aprendizaje competitivo y el aprendizaje por corrección de error. No es la intención de en este artículo abordar con lujo de detalles cada uno de los algoritmos de aprendizaje. Por el contrario, sólo serán descritos brevemente algunos de ellos con el objeto de aclarar conceptos y fijar algunas ideas básicas.

El aprendizaje Hebbiano (Hebb, 1949) y el aprendizaje competitivo (Willshaw y von der Malsburg, 1976) son reglas basadas en consideraciones neurobiológicas.

En el caso particular del aprendizaje Hebbiano, tal y como se comentó en la primera sección de este artículo, las sinapsis neuronales (pesos) se refuerzan o debilitan según sea la historia de su actividad. De acuerdo a este postulado, la actualización de los pesos se hace de la siguiente manera:

$$\Delta W_{ij}(n) = \eta x_i(n) y_j(n) \quad \text{Ecuación 6}$$

donde x_i y y_j constituyen la actividad pre-sináptica y post-sináptica (entrada y salida de la neurona en cuestión) respectivamente, y η se denomina la constante de aprendizaje. Versiones más elaboradas de esta regla de aprendizaje consideran la inclusión de un término de *olvido*, que evita el posible crecimiento exponencial de los pesos durante el entrenamiento de la red (Kohonen, 1988).

En el aprendizaje competitivo, las neuronas de una red compiten por activarse (responder) ante la presencia de un estímulo dado. Generalmente, a diferencia de otros tipos de entrenamiento, en el aprendizaje competitivo sólo una neurona se activa a la vez. En la regla de aprendizaje competitivo, sólo se actualizan los pesos de la neurona que responde con más fuerza a un estímulo dado. Según este postulado, la actualización de los pesos se hace de la siguiente manera:

$$\Delta W_{ij}(n) = \eta [x_i(n) - W_{ij}(n)] \delta(k-j) \quad \text{Ecuación 7}$$

donde x_i representa el estímulo en un momento dado, k es el índice que identifica a la neurona ganadora y $\delta(k-j)$ es una función pulso que vale 1 cuando $k=j$ y 0 cuando $k \neq j$. Como resultado del uso de una regla de aprendizaje competitivo, cada una de las neuronas de la red termina por asociarse a un estímulo o grupo de estímulos en particular, por lo cual, las redes que operan bajo esta regla de entrenamiento son ideales para abordar problemas de clasificación basados en análisis de clusters (Haykin, 1994).

A diferencia de los antes expuestos, el aprendizaje por corrección de error tiene sus fundamentos en la teoría de filtraje óptimo, la cual fue brevemente discutida en la segunda sección de este artículo. Según esta regla de aprendizaje, también conocida como la regla Delta (Widrow and Hoff, 1960), la actualización de los pesos se hace de acuerdo a la siguiente expresión:

$$\Delta W_{ij}(n) = \eta x_i(n) e_j(n) \quad \text{Ecuación 8}$$

donde x_i representa el estímulo y e_j es el error de ajuste tal y como se definió en la ecuación 2.

Otro aspecto determinante en el proceso de entrenamiento de una red neuronal es el denominado paradigma de aprendizaje. El paradigma de aprendizaje depende de la forma en que la red neuronal interactúa con su entorno, lo cual depende directamente de la naturaleza del problema en cuestión y los datos disponibles.

Existen tres paradigmas de aprendizaje en inteligencia artificial; ellos son: aprendizaje supervisado, aprendizaje por refuerzo y aprendizaje no supervisado.

En el aprendizaje supervisado existe la figura de un “tutor” o “maestro”, externo a la red neuronal, capaz de cuantificar el desempeño de la red. En el contexto de las aplicaciones prácticas, esto se traduce en la disponibilidad de un subconjunto de datos de entrada o estímulos para los cuales se conoce su salida o respuesta. Este subconjunto de datos es el denominado conjunto de entrenamiento. El ejemplo clásico de aprendizaje supervisado lo constituyen los perceptrones de múltiples capas entrenados mediante el algoritmo de *backpropagation* (propagación hacia atrás del error). *Backpropagation* es una versión generalizada del problema de minimización del error cuadrático medio (el cual fue definido en la ecuación 3); y debe su nombre a la forma en que el gradiente del error respecto a los pesos de las neuronas en cada capa es calculado y propagado hacia “atrás” desde la capa de salida hacia las capas previas (Haykin, 1994).

El aprendizaje por refuerzo, debe su nombre al hecho de que existe la figura de un “crítico”, externo a la red neuronal, que es capaz de decidir si el desempeño de la red fue satisfactorio o no satisfactorio, pero no tiene manera alguna de cuantificar el desempeño de la red (Barto, 1992). De esta forma, cuando el desempeño es satisfactorio, el “crítico” “recompensa” a la red y la tendencia de la misma para responder al estímulo de esa forma particular es reforzada. Por el contrario, si el desempeño no es satisfactorio la tendencia a responder de esa forma es debilitada. A diferencia del aprendizaje supervisado, en el cual se conoce un conjunto de respuestas deseadas a partir de las cuales el sistema es entrenado, en el aprendizaje por refuerzo, el sistema debe “explorar” por sí mismo el dominio de los datos y aprender basado en las penalizaciones o recompensas que recibe.

Finalmente, en el aprendizaje no supervisado, no existe ningún agente externo a la red neuronal capaz de guiar el proceso de aprendizaje (Becker, 1991). En este caso el aprendizaje se lleva a cabo mediante la extracción por parte de la misma red neuronal de las tendencias estadísticas existentes en los datos. De esta manera la red neuronal se convierte en una representación o mapa del espacio de los datos. Este paradigma de aprendizaje es el más comúnmente usado en los problemas de clasificación basados en análisis de clusters y reconocimiento de patrones.

Redes neuronales en caracterización de yacimientos

En los últimos años ha proliferado el uso de redes neuronales en problemas de estimación y clasificación en caracterización de yacimientos. A la fecha, existen gran cantidad de programas comerciales que permiten estimar y/o identificar propiedades de yacimientos, tipo de litología, ambientes sedimentarios, a partir de datos de campo usando diversas arquitecturas de redes neuronales y distintos paradigmas de entrenamiento. Entre los más populares podemos mencionar:

- *SEISCLASS™*. Es una herramienta de la compañía Schlumberger para clasificación de facies sísmicas que ofrece, entre otros métodos, redes neuronales no supervisadas como una alternativa para realizar análisis de clusters.
- *ROCKCELL™*. También de Schlumberger, ofrece la posibilidad de usar redes supervisadas y no supervisadas para la identificación de electro-facies a partir de registros de pozo con posible calibración o no con datos de núcleo.
- *EMERGE™*. Es una herramienta de la compañía Hampson & Russell, que permite predecir parámetros del yacimiento usando combinación de múltiples atributos sísmicos. Este programa, basado en el uso de esquemas supervisados, ofrece las alternativas de perceptrones de múltiples capas y redes neuronales probabilísticas (Hampson *et al.*, 2001).
- *GDI™*. Es una herramienta de la compañía dGB. Este programa ofrece tanto técnicas supervisadas, como no supervisadas. Las técnicas supervisadas son usadas para problemas de clasificación, cómo identificación litológica, y para estimación de propiedades. Las técnicas no supervisadas son usadas para análisis de clusters en aplicaciones cómo por ejemplo, la clasificación de formas de ondícula sísmica a lo largo de un horizonte.
- *D-TECT™*. También de dGB, es una herramienta de cálculo de atributos y reconocimiento de patrones que permite entre otras cosas clasificación de facies sísmicas mediante esquemas supervisados.
- *LITHANN™*, es la herramienta de la compañía Rock Solid Images para clasificación y calibración de volúmenes de múltiples atributos. Este programa utiliza un esquema de clasificación no supervisada basado en el uso de redes neuronales de Kohonen (K-SOM: Kohonen's Self Organizing Map). (Kohonen, 1990)

En Venezuela, también se han desarrollado numerosas aplicaciones basadas en el uso de redes neuronales para problemas de estimación y clasificación. Cómo ejemplo de ellas se pueden mencionar:

- Clasificación con redes neuronales de rangos de porosidad y permeabilidad a partir de datos de resonancia magnética nuclear medidos en núcleos (Rodrigues *et al.*, 2000).
- Análisis de similitud sísmica basado en redes neuronales de Kohonen (Rondón y Banchs, 2002) y de Hopfield (Banchs y Jiménez, 2002).
- Estimación de cubos de pseudo registros y/o propiedades de yacimientos con sus respectivos intervalos de confianza usando perceptrones de capas múltiples (Banchs y Michelena, 2000, 2002).

- Combinación de métodos geoestadísticos y redes neuronales para mejorar la calidad de las estimaciones de mapas de propiedades a partir de datos de pozos (Blunda *et al.*, 2002).

En las siguientes dos secciones se muestran dos ejemplos específicos de la aplicación de algunas de estas técnicas a datos de campo en yacimientos venezolanos. En el primer caso se presenta un ejemplo de clasificación litológica, y en el segundo se presenta un ejemplo de estimación de pseudo-propiedades de yacimiento.

Clasificación litológica

En esta sección se ilustra un ejemplo de clasificación litológica usando los datos sísmicos de un campo al sureste de Venezuela. En este ejercicio se realizó una clasificación de tipos de arenas a partir de los resultados de un análisis fractal basado en wavelets de los datos sísmicos en un horizonte de interés. El análisis fractal basado en wavelets (WBFA) es una técnica que permite analizar cómo varía la energía de la señal respecto a la escala (Jiménez *et al.*, 1999).

En el área donde se realizó el ejercicio, se disponía de información de 24 registros de pozos, una caracterización a priori de los ambientes sedimentarios en un intervalo de interés, un volumen de datos sísmicos y los horizontes interpretados en el intervalo de interés. El objeto de este estudio era evaluar la factibilidad de delimitar las áreas que presentan mayor contenido de arenas limpias, las que presentan arenas sucias y sus respectivas zonas de transición. El método se aplicó a los pozos en primer lugar, utilizando registros del tipo gamma ray, donde se logró apreciar una evidente separación de los distintos tipos de arena. Para obtener un mapa de toda el área se aplicó el método WBFA al volumen de datos sísmicos y se entrenó una red neuronal con la información correspondiente a las localidades de los pozos.

Para este ejemplo se usó un perceptrón de capas múltiples y el algoritmo de aprendizaje de *backpropagation*. En la figura 5 se muestra el mapa de distribución de arenas obtenido para el intervalo de interés. En dicho mapa el color amarillo representa las arenas de buena calidad, el color gris representa las arenas de mala calidad y el rojo las zonas de transición.

El porcentaje de éxito estimado para la clasificación fue de aproximadamente un 80%, el cual se puede verificar a partir de las clases obtenidas para cada uno de los pozos usados en el estudio. Adicionalmente, el resultado obtenido corrobora el modelo sedimentológico previamente existente para el área, según el cual la cantidad de arenas de mala calidad aumenta en la dirección sudoeste.

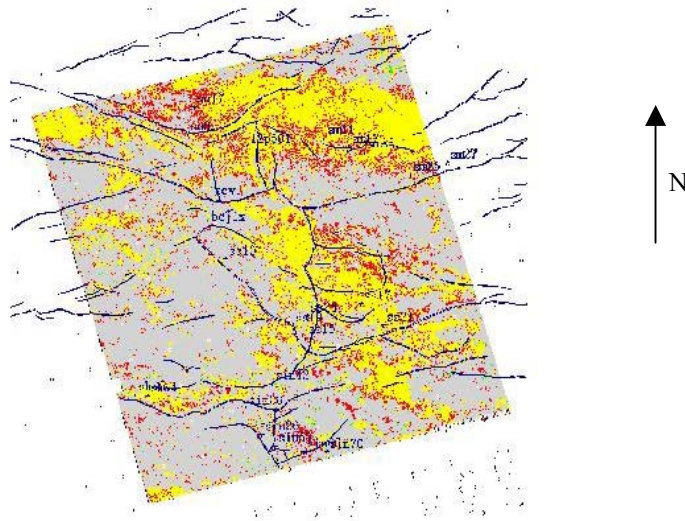


Figura 5: Mapa de distribución de arenas.

Volumen de pseudo registros

En este ejemplo se estima un volumen de potencial espontáneo a partir de un volumen de atributos sísmicos 3D. En esta área en particular el potencial espontáneo resulta ser un buen indicador litológico. La extensión total del área de interés es de 5.73 Km² y el intervalo de interés es de 100 ms. Un total de 25 registros de potencial espontáneo fue usado para entrenar la red neuronal, la cual fue un perceptrón de capas múltiples constituido por 4 capas (con 10, 12, 12 y 1 neuronas en cada capa, respectivamente). La red se entrenó con el algoritmo de *backpropagation*.

En este ejercicio se llevó a cabo una estimación de intervalos de confianza para el volumen generado (Banchs y Michelena, 2000). Para tal fin, un total de 70 estimaciones independientes fueron realizadas. De esta forma, se obtuvo un total 70 volúmenes de potencial espontáneo, para los cuales se observaron sus distribuciones y se calcularon sus estadísticos. Asumiendo una ley de distribución gaussiana, se generó el volumen de potencial espontáneo más probable a partir de los valores medios de la propiedad en cada punto del volumen. Para calcular el volumen de intervalos de confianza, se utilizó la varianza de la propiedad en cada punto del volumen. La figura 6 muestra una sección del volumen de potencial espontáneo y su respectiva medida de incertidumbre.

En la sección promedio de potencial espontáneo mostrada en la figura los valores más negativos del potencial espontáneo (que se interpretan como arenas) se ilustran en azul y los valores menos negativos (que se interpretan como arcillas) se ilustran en rojo. Obsérvese en la figura la presencia de “lentes” de arena; este resultado concuerda con lo establecido en el modelo sedimentológico de este yacimiento.

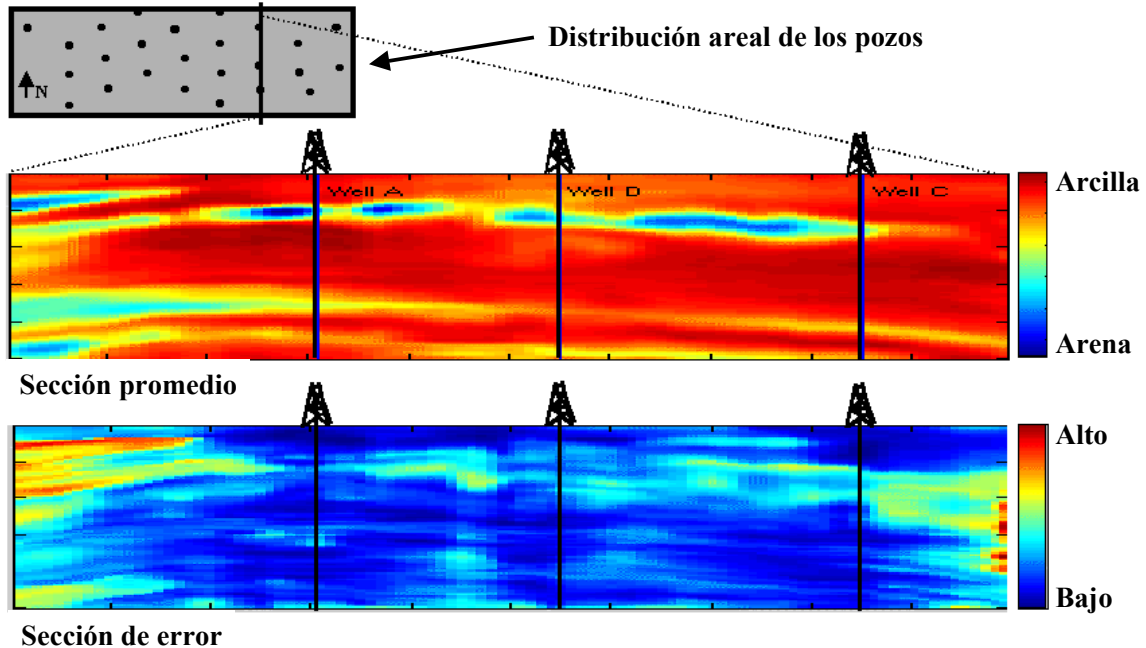


Figura 6: Secciones estimadas de potencial espontáneo y margen de error.

Conclusiones y recomendaciones

El uso de redes neuronales en problemas de estimación y clasificación en caracterización de yacimientos ha incrementado en forma espectacular en los últimos años. Sin embargo, a estas alturas todavía las redes neuronales son vistas por muchos como cajas negras y misteriosas capaces de dar respuestas que muchas veces no se sabe como validar o verificar.

En el fondo, las redes neuronales no son más que filtros no lineales que tienen la gran ventaja de poder ser “entrenados” para representar relaciones complejas entre conjuntos de datos, que los métodos lineales tradicionales de interpolación y/o clustering no son capaces de representar. Este trabajo pretendió desmitificar un poco el concepto de las redes neuronales mediante la presentación de los aspectos más fundamentales de su teoría; e ilustrar su uso en problemas específicos de caracterización de yacimientos.

Aunque las redes neuronales son y seguirán siendo una gran herramienta de trabajo en problemas de inferencia, deben usarse con mucho cuidado. El principal problema de las redes neuronales es que siempre dan una respuesta. Sin embargo, debe tenerse en cuenta que la calidad y validez de una representación obtenida a partir de una red neuronal dependerá siempre de algunos aspectos prácticos que la mayoría de las veces no son sencillos de controlar. Entre ellos podemos mencionar el hecho de que tan bien los datos de entrenamiento representan el espacio de los datos, ó que tan adecuada es una arquitectura

específica para un problema en cuestión, así como cuál es el número de neuronas óptimo en la red para resolver un problema específico, ó cual es el conjunto de variables de entrada que realmente son significativos para la solución de un problema. Son precisamente estos problemas, para los cuales todavía no hay respuestas teóricas claras, los que le dan esa imagen algo mítica y misteriosa al uso de estas herramientas.

En definitiva, se puede decir que el uso de redes neuronales en cualquier contexto es en parte una técnica y en parte un arte. Y es precisamente de la experiencia que da su propio uso que se refina el técnico y se perfecciona el artista.

Referencias

Banchs, R.; Jiménez, J.; 2002, "Content addressable memories for seismic attribute pattern recognition", *64th EAGE Conference & Exhibition*.

Banchs, R.; Michelena, R.; 2000, "Well log estimates and confidence intervals by using artificial neural networks", *SEG 2000 expanded abstracts*.

Banchs, R.; Michelena, R.; 2002, "From 3D seismic attributes to pseudo well log volumes using neural networks: practical considerations", *The Leading Edge*, octubre.

Barto, A.; 1992, "Reinforcement learning and adaptive critic methods", *Handbook of Intelligent Control*, 469-491, New York, Van Nostrand-Reinhold.

Becker, S.; 1991, "Unsupervised learning procedures for neural networks", *International Journal of Neural Systems*, 2, 17-33.

Blunda, Y.; Camacho, L.; Porjesz, R.; Banchs, R.; 2002, "Estimaciones de porosidad empleando métodos basados en geoestadística y redes neuronales", *XI Congreso Venezolano de Geofísica*.

Fahlman, S.; Hinton, G.; 1987, "Connectionist architectures for artificial intelligence", *IEEE Computer*, **20**, 100-109.

Hampson, D.; Schuelke, J.; Quirein, J.; 2001, "Use of multiattribute transforms to predict log properties from seismic data", *Geophysics*, **66**, 220-236.

Haykin, S.; 1994, *Neural networks: a comprehensive foundation*, New York, Macmillan.

Haykin, S.; 1996, *Adaptive filter theory*, Upper Saddle River, Prentice-Hall.

Hebb, D.; 1949, *The organization of behavior: a neuropsychological theory*, New York, Wiley.

Holland, J.; 1992, *Adaptation in natural and artificial systems*, Cambridge, MIT Press.

Jimenez, J.; Michelena, R.; Peinado, A.; 1999, "Facies recognition using wavelet-based fractal analysis on compressed seismic data", *SEG 1999 expanded abstracts*.

Kohonen, T.; 1988, "An introduction to neural computing", *Neural Networks*, **1**, 3-16.

Kohonen, T.; 1990, "The self-organizing map", *Proceedings of the IEEE*, **78**, 1464-1480.

McCulloch, W.; Pitts, W.; 1943, "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics*, **5**, 115-133.

McNeill, D.; Freiberger, P.; 1993, *Fuzzy Logic*, New York, Simon & Schuster.

Oppenheim, A.; Schafer, R.; 1989, *Discrete-time signal processing*, Englewood Cliffs, Prentice-Hall.

Rodrigues, P.; Alcocer, Y.; Matheus, J.; 2000, "NMR Signal Pattern Classification for Estimation of Petrophysical Properties", *SPE Annual Technical Conference and Exhibition*. SPE 63258

Rondón, O.; Banchs, R.; 2002, "Obtención de mapas de similitud sísmica 2D mediante redes neuronales", *XI Congreso Venezolano de Geofísica*.

Waterman, D.; 1986, *A guide to expert systems*, Addison-Wesley.

Widrow, B.; Hoff, M.; 1960, "Adaptive switching circuits", *IRE WESCON Convention record*, 96-104.

Widrow, B.; Stearns, S.; 1985, *Adaptive signal processing*, Englewood Cliffs, Prentice-Hall.

Willshaw, D.; von der Malsburg, C.; 1976, "How patterned neural connections can be set up by self-organization", *Proceedings of the Royal Society of London*, **194**, 431-445.