

Redes Neuronales Aplicadas al Análisis de Datos

Juana Jenny Velasco Avendaño
Simulación de Sistemas (INF-391)
Jenavela@gmail.com

RESUMEN

En este documento se trata de entender el proceso que realizan las redes neuronales artificiales al analizar datos ya que actualmente las RNA son ampliamente utilizadas en la generación de modelos de clasificación, constatándose en diversos estudios su superioridad frente a otras técnicas de clasificación además se determinó que las RNA son más eficaces que los algoritmos de inducción de reglas y el análisis discriminante, permiten apoyar la toma de decisiones en prestigiosas compañías alrededor del mundo, a través de un proceso de entrenamiento, las RNA, son capaces de determinar relaciones no lineales entre un conjunto de datos, asociando a patrones de entrada la o las salidas correspondientes. De ahí que la aplicación de las RNA en el ámbito de la Ingeniería de Tráfico puede ayudar a cumplir su objetivo global; agentes/sistemas "inteligentes", capaces de adaptar los diferentes flujos de datos de acuerdo a los recursos disponibles.

Palabras clave

Redes neuronales artificiales, Análisis de datos, proceso, entrada, salida.

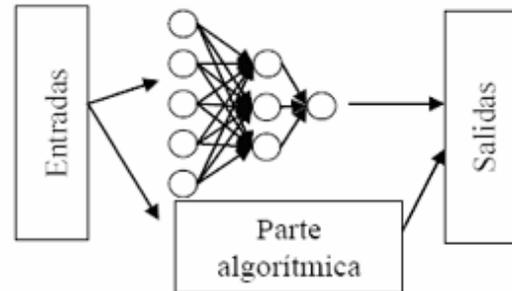
1. INTRODUCCION

Las Redes Neuronales Artificiales (Artificial Neural Networks, ANN) son unas estructuras computacionales que han sido inspiradas en sistemas neuronales biológicos. Básicamente, una ANN está conformada por nodos llamados neuronas y por las conexiones existentes entre estos nodos. Su objetivo principal es transformar un conjunto de datos de entrada en un conjunto de datos de salida en donde entradas y salidas son datos numéricos. Esta transformación se da a través de la interconexión de las neuronas que conducen los datos desde una capa de entrada hasta una capa de salida.

De manera formal Rumelhart definió los componentes

De un sistema neuronal:

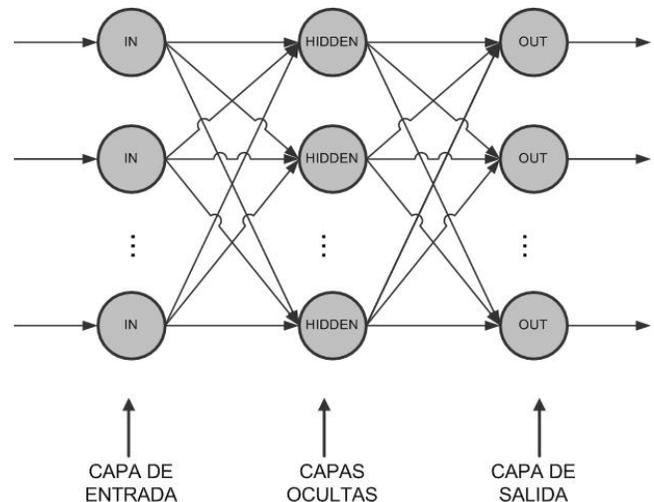
1. Un conjunto de procesadores elementales o neuronas
2. Un patrón de conectividad entre las neuronas o arquitectura
3. Una dinámica de activaciones.
4. Una regla de aprendizaje
5. El entorno donde opera



Sistema Neuronal Artificial

2. MODELO GENERAL DE UNA RED NEURONAL

La siguiente figura se muestra; la capa de entrada recibe todo el conjunto de variables numéricas, se puede contar con cero o más capas ocultas y una capa de salida donde se encuentran las variables dependientes. Como recomendación, los datos de entrada deben ser numéricos y debe estar normalizado. Adicionalmente, las conexiones entre las neuronas juegan un papel importante en el sistema, pues en ellas se determina cómo debe funcionar el modelo; esto significa que cada conexión posee un peso lo cual influye en las capacidades de clasificación y/o predicción.



De acuerdo a algunos autores el número de capas ocultas y el número de neuronas en las capas ocultas es uno de los problemas más difíciles de tratar; adicionalmente, los autores estableces que con una sola capa oculta es suficiente para modelar un sistema complejo que produzca una buena exactitud.

Por otra parte, establecen que el número de neuronas en las capas ocultas y el número de capas ocultas puede ser determinado a través de ensayo y error.

3. FUNCIONAMIENTO UNA RED NEURONAL

Para poner en funcionamiento se deben considerar tres fases importantes:

- 1) Construcción de la arquitectura de ANN. Aquí se determinan las variables de entrada, la clase o variable dependiente a predecir y la topología de la red. Tanto las variables de entrada como de salida deben ser previamente normalizados para evitar que algunos atributos tengan más influencia que otros...
- 2) Entrenamiento y validación del modelo. En esta fase se ajustan los pesos de las conexiones hasta que el error resultante del modelo sea lo más reducido posible, i.e. la diferencia entre la salida esperada y la real tienda a ser cero.
- 3) Predicción y/o clasificación. Una vez entrenado el modelo, se procede a mostrarle a la ANN nuevos casos. Estos casos no han sido tratados anteriormente por el sistema. El modelo de ANN debe estar en capacidad de clasificar y/o predecir cada nuevo caso. Sin embargo, se ha considerado 7 etapas para la construcción

Y puesta en funcionamiento de la ANN, pero para efectos prácticos las tres fases anteriores expresan la misma idea. Para el caso de este trabajo y apoyándose en [19], una red neuronal adecuada en la construcción de un modelo de smoking puede

Estar conformada por:

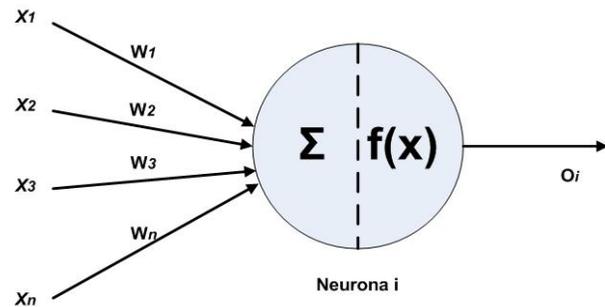
- _ Tantas neuronas en la capa de entrada como atributos numéricos posea el individuo solicitante del crédito.
- _ Una neurona en la capa de salida que clasifique el comportamiento crediticio del cliente; usualmente, bueno o malo. Aunque se podrían considerar más neuronas en esta capa dependiendo de las necesidades que se deseen cubrir.
- _ Una capa intermedia u oculta.

Una vez definida la arquitectura o la topología de la Red Neuronal, es necesario iniciar el proceso de aprendizaje o entrenamiento. En este proceso, como se mencionó anteriormente, se desea encontrar la combinación adecuada de pesos en las conexiones que conlleven a la reducción del error en la capa de salida. La Figura 4 muestra una neurona de una capa oculta de una Red Neuronal (e.g. Backpropagation Backpropagation Neural Network, BPN).

De manera general, cada neurona en determinada capa toma el conjunto de pesos multiplicado por la entrada correspondiente en donde inicialmente los pesos son asignados de manera aleatoria. Esto es:

$$I = \sum w_i x_i + \theta$$

Donde I es la entrada de la neurona y θ es un peso adicional que influencia el resultado (i.e. bias). Posteriormente se procesa la entrada a través de una función de activación, i.e. la neurona se excita o se inhibe dependiendo del valor I. Existen



Varios tipos de funciones de activación que suelen producir valores entre 0 y 1 o entre -1 y 1, i.e. el resultado se encuentra normalizado. Algunas de las funciones de activación $f(I)$ más empleadas son la función sigmoideal, la función logística y la función tangente hiperbólica. En la mayoría de los casos, las funciones de activación a emplear en la red neuronal son no lineales (i.e. salida de la neurona).

Finalmente, los pesos deben ser ajustados iterativamente de acuerdo a un criterio apropiado como por ejemplo la minimización del Error Cuadrático Medio (Sum of Squared Error, SSE):

$$SSE(W) = \frac{1}{2} \sum (y_i - \hat{y}_i)^2$$

Donde el SSE depende de W debido a que la clase estimada \hat{y} es una función de los pesos asignada a las neuronas de las capas ocultas y la capa de salida (i.e. $f(I) = \hat{y}$). Para apoyar el proceso de aprendizaje se utiliza la tasa de aprendizaje (learning rate). La decisión de la escogencia de la tasa de aprendizaje es crucial, pues si se escoge una tasa muy pequeña, el proceso de aprendizaje se vuelve lento y la convergencia es demorada. Por otra parte, si la tasa es grande provocará una oscilación en el modelo y la convergencia no se dará. Para el caso de una BPN, uno de los algoritmos de aprendizaje más empleados es el Gradiente Descendiente.

4. TASA DE APRENDIZAJE

La tasa de aprendizaje de la RNA depende de distintos factores controlables que hay que tener en cuenta. Obviamente, un valor bajo de tasa de entrenamiento equivale a que necesitamos gastar mucho tiempo para llevar a cabo el entrenamiento para producir una RNA bien entrenada. Con valores de entrenamiento mayores, la red puede no ser capaz de discriminar de manera tan adecuada como un sistema que aprende más despacio. Generalmente, factores adicionales a parte del tiempo tienen que ser considerados cuando se discute el entrenamiento off-line:

- Complejidad de la red: tamaño, paradigma, arquitectura
- Tipo de algoritmo de aprendizaje empleado.
- El error admitido de la red final.

Si cambiamos alguno de estos factores podemos o aumentar el tiempo de entrenamiento a un valor elevado u obtener un error inaceptable. La mayoría de algoritmos de aprendizaje requieren que se les pase una tasa de aprendizaje, o un término una constante de aprendizaje. Normalmente este término es un valor positivo entre 0 y 1. Si la tasa de aprendizaje es mayor que 1, es fácil para el algoritmo de aprendizaje provocar que se creen

pesos incorrectos y la red oscilaría. Pequeños valores de tasa de aprendizaje no corregirán el error actual de manera rápida.

Equivalentes artificiales de los dispositivos biológicos

Aquí se muestra la equivalencia entre los elementos artificiales y los biológicos a los fines de reproducir la trama neuronal

ELEMENTO BIOLÓGICO	ELEMENTO ARTIFICIAL
Neurona	<ul style="list-style-type: none"> • NODO DE LA RED = POSICIÓN DE MEMORIA RAM + ESPACIO DISPONIBLE DE DISCO RÍGIDO + PROGRAMA DE SOFTWARE NEURONAL
Entradas Sensoriales	<ul style="list-style-type: none"> • DIGITALIZACIÓN DE IMÁGENES • ARCHIVOS DE COMPUTADORA DE DIVERSOS FORMATOS
Salidas Sensoriales Comportamiento	<ul style="list-style-type: none"> • CREACIÓN DE ARCHIVO DE COMPUTADORA CON DATOS PROCESADOS • ACTIVACIÓN DE PROGRAMA DE COMPUTADORA PREDEFINIDO • ACTIVACIÓN
Interconexión Sináptica	<ul style="list-style-type: none"> • PREFERENCIAS Y PONDERACIONES DE LA TRANSFERENCIA DE VALORES ENTRE LAS DISTINTAS POSICIONES DE MEMORIA

SENSITIVITY NEURAL NETWORK 1.0: UN SIMULADOR DE REDES NEURONALES ARTIFICIALES CON ANÁLISIS DE SENSIBILIDAD

El programa informático *Sensitivity Neural Network 1.0* (SNN) permite simular el comportamiento de una arquitectura neuronal del tipo perceptrón multicapa (con una capa de entrada, una capa oculta y una capa de salida) entrenada mediante la regla de aprendizaje *backpropagation error* e incorpora como novedad diversos métodos --denominados genéricamente análisis de

sensibilidad--, encaminados al estudio del efecto de cada variable de entrada sobre la salida de la red. El SNN ha sido creado mediante el lenguaje de programación Borland Delphi 5 y desarrollado para el entorno Windows, por lo que su interfaz de usuario es muy amigable. El interfaz consiste en una ventana única dividida en una serie de secciones que permiten de forma muy sencilla la manipulación de los aspectos más relevantes del entrenamiento y validación de una red perceptrón multicapa. La sección *Datos* permite visualizar las matrices de datos utilizadas para el entrenamiento, validación y test del modelo neuronal. En la sección *Parámetros de aprendizaje* se puede determinar el valor de la tasa de aprendizaje y el factor momento. La sección *Arquitectura de la red* permite configurar el número de neuronas en la capa oculta y el tipo de función de activación (lineal, sigmoidea logística y sigmoidea tangente hiperbólica) de las neuronas de la capa oculta y de salida. En la sección *Criterios de parada* se determina el momento en el que el entrenamiento de la red finaliza. La sección *Pesos* permite determinar los pesos iniciales de conexión y de umbral de la red mediante un valor semilla o importar los pesos obtenidos mediante la utilización de otro programa simulador. En la sección *Estadísticas* se proporcionan una serie de índices y representaciones gráficas que describen el rendimiento del modelo neuronal durante el proceso de entrenamiento. Una vez obtenido el modelo final, se puede obtener un informe en donde se recoge la salida o predicción de la red para un conjunto de patrones dados, diversos índices globales de rendimiento, las matrices de pesos de la red obtenida y el resultado de la aplicación de los análisis de sensibilidad más ampliamente descritos en la literatura de redes neuronales artificiales.

5. CONCLUSIONES

Dando su flexibilidad y capacidad de generalización como su habilidad para diagnosticar con notable precisión en diversas áreas, permiten afirmar que las RNAs se convertirán en un futuro próximo en una herramienta difundida en el análisis de problemas biomédicos. Sin embargo, y a pesar de sus éxitos iniciales, se requiere aún una considerable cantidad de trabajo antes que las RNAs puedan ser aceptadas como una asistencia clínica legítima...

6. RECONOCIMIENTOS

Agradecimiento al Lic. Aldo Valdez por darnos la iniciativa a la investigación.

7. REFERENCIAS

- [23] Pacific Northwest National Laboratory
<http://www.emsl.pnl.gov:2080/proj/neuron/>
- [24] www.thesisenxarxa.net
- [25] www.invenia.es
- [26] dialnet.unirioja.es/servlet
- [27] http://es.wikipedia.org/wiki/Red_neuronal_artificial