

Resumen

Las redes neuronales artificiales han demostrado en estos últimos años su valía a la hora de resolver problemas complejos que con la programación tradicional son difíciles de resolver. Por ejemplo, el código necesario para programar la operación de reconocer el rostro de una persona es muy complejo, pero resulta una tarea relativamente sencillo para el cerebro humano, y relativamente fácil de implementar en una red neuronal. La inteligencia artificial microscópica parte de la estructura que presenta el cerebro produciendo modelos influenciados por dicha estructura. El resultado son las Redes Neuronales Artificiales (RNA) que están inspiradas en las redes neuronales biológicas del cerebro humano, y han sido propuestas como modelos extremadamente simplificados de su funcionamiento presentando resultados razonables en aplicaciones donde las entradas presentan ruido o están incompletas, tales como conversión texto a voz, procesado natural del lenguaje, compresión de imágenes, reconocimiento de patrones en imágenes, problemas de combinatoria, procesado de señales, etc.

La neurona es la unidad fundamental del sistema nervioso y en particular del cerebro que consta de un gran número de ellas y están altamente interconectadas. La potencia computacional del cerebro deriva, principalmente, de su estructura de cálculo distribuido paralelo, y aunque las RNA sean un modelo muy esquemático, presentan una riqueza sorprendente de estados y de comportamientos. Una red neuronal consiste en un conjunto de neuronas artificiales conectadas de una forma concreta. Existen diferentes tipos de redes neuronales. Una red típica conocida como *feedforward* o de procesamiento hacia adelante, consiste en una secuencia de capas con conexiones entre capas adyacentes consecutivas.

El aprendizaje con redes neuronales artificiales de tipo *feedforward* se realiza principalmente siguiendo dos enfoques: 1) Algoritmos derivados de la técnica de retropropagación de errores (*Backpropagation*), los cuales necesitan de la especificación a priori del número de unidades ocultas junto con sus conexiones, que utilizan la minimización de una función de costo para obtener los pesos sinápticos. 2) Enfoques constructivos, en los cuales al mismo tiempo que se aprende el conjunto de entrenamiento, se genera una arquitectura de red apropiada. Las redes neuronales constructivas (CoNN) son por lo tanto una colección de algoritmos que alteran la estructura de la red en el aprendizaje produciendo automáticamente una red con un tamaño adecuado para un problema concreto adaptándose en cada momento al problema que se quiere aprender.

La presente tesis doctoral se enmarca dentro del área de las redes neuronales cons-

tructivas, siendo el principal objetivo proponer, desarrollar y evaluar nuevos algoritmos, analizando su aplicación a problemas reales del ámbito de la ingeniería. Con este objetivo se han desarrollado dos nuevos algoritmos constructivos: DASG y C-Mantec, los cuales han sido aplicados con éxito a diversos tipos de problemas en las áreas de síntesis de arquitecturas y de clasificación de patrones.

DASG (*Decomposition Algorithm for Synthesis and Generalization*) es un algoritmo de redes neuronales constructivo basado en un enfoque secuencial, siendo capaz de clasificar y sintetizar conjuntos de datos binarios con salida biclase. DASG es un algoritmo de descomposición de funciones booleanas que se basa principalmente en la descomposición de la función objetivo en dos nuevas funciones de menor complejidad. Aplicando iterativamente este proceso, el algoritmo consigue obtener un conjunto donde todas las funciones son linealmente separables. DASG genera redes similares a las redes *feedforward*, compuestas por una capa de entrada, una única capa oculta y una capa de salida. En la capa de salida el algoritmo siempre tendrá implementado el operador lógico AND o el operador lógico OR de acuerdo a las características de la función objetivo. Además, DASG tiene la ventaja de que puede extraerse fácilmente el conocimiento incorporado, es decir, una vez entrenada la red se pueden extraer reglas potencialmente interpretables de manera directa.

C-Mantec (*Competitive MAjority Network Training Error Correcting*) es un algoritmo de tipo constructivo que genera arquitecturas de redes neuronales muy compactas con buena capacidad de generalización. Utiliza la competición entre las neuronas en la fase de entrenamiento, evitando la congelación de los pesos sinápticos y permitiendo que todas las neuronas puedan aprender en cualquier etapa del procedimiento, un elemento nuevo dentro de las características usuales de los algoritmos constructivos. Este enfoque cooperativo del proceso de aprendizaje, conjuntamente con la incorporación de un proceso de filtrado de ejemplos para evitar el efecto de sobreentrenamiento, permiten que el algoritmo consiga una muy buena capacidad de predicción.

Ya en el campo de las aplicaciones, se ha aplicado primero el algoritmo DASG al problema de predicción de recidiva de cáncer de mama, realizándose el proceso de extracción de conocimiento a través de la obtención de las reglas que genera el algoritmo. En relación al algoritmo C-Mantec se muestran diversas aplicaciones dentro del área de redes de sensores, presentándose primero la aplicación a tres problemas diferentes, para luego analizar en detalle un caso de estudio. El estudio de esta aplicación se centró en comparar la eficiencia energética relacionada con la reprogramación dinámica de los nodos de una red de sensores cuando se utiliza el algoritmo C-Mantec, y compararla con el uso de otros métodos tradicionales.