

REDES DE NEURONAS ARTIFICIALES

Curso 2017-18

PRÁCTICA I. PROBLEMA DE REGRESIÓN

Resistencia a la compresión del hormigón

Introducción

El objetivo de esta práctica es abordar un problema real de regresión utilizando dos modelos de redes de neuronas supervisados:

- El modelo lineal Adaline
- El modelo no-lineal Perceptron Multicapa

Se trabajará con el dominio real de regresión conocido como “Concrete Compressive Strength (Resistencia a la compresión del hormigón)”. El hormigón es uno de los materiales más utilizados en la ingeniería civil. Para los expertos en este campo es muy importante conocer la resistencia de dicho material a la fuerza de compresión. Además, es de gran interés conocer dicha resistencia en función de variables como la edad del hormigón o los ingredientes que lo componen, evitando así el tener que realizar ensayos con el propio material antes de su uso.

Para abordar el problema se dispone de un conjunto de datos que contiene 1030 instancias con 8 variables de entrada: la edad del hormigón en días y las cantidades de 7 componentes (cemento, escoria de alto horno, cenizas volantes, agua, superplastificante, agregado grueso y agregado fino), medidas en kg por m³ de mezcla. La variable de salida o variable a predecir es la resistencia que se observó en el hormigón a la fuerza de compresión, medida en MPa (megapascuales). En el siguiente enlace <http://sci2s.ugr.es/keel/dataset.php?cod=44> puede encontrarse una breve descripción del conjunto de datos.

Los datos que serán utilizados para la realización de la práctica podrán descargarse en la sección Practica I en Aula Global.

Trabajo a realizar

1. Preparación de datos

Antes de realizar el aprendizaje de las redes, hay que realizar un preproceso de los datos disponibles:

- **Normalización:** Es recomendable normalizar las variables de entrada y salida en el rango [0, 1]. Para la normalización se calculará el valor mínimo y máximo de cada variable i y se aplicará la siguiente transformación lineal:

$$VarNorm_i = (VarOrig_i - ValorMin_i) / (ValorMax_i - ValorMin_i)$$

- **Aleatorización:** Para que el entrenamiento de las redes se realice en condiciones adecuadas es importante desordenar o ‘aleatorizar’ los datos.
- **Separación de tres conjuntos de datos:**
 - Datos de **entrenamiento (70% del total de datos)** para realizar el aprendizaje de la red.
 - Datos de **validación (15% del total de datos)** que serán utilizados para elegir los valores óptimos de los parámetros de la red (razón de aprendizaje, número de ciclos, número de neuronas).

- Datos de **test (15% del total de datos)** para evaluar la capacidad de generalización de la red.

2. Desarrollo y experimentación con Adaline

Debido a la sencillez de esta red, no se utilizará un simulador sino que se desarrollará un programa que realice el aprendizaje del Adaline, explicado en el Tema 2. Dicho programa se puede desarrollar en el lenguaje de programación deseado.

Para verificar que el programa funciona correctamente, es decir, que simula el aprendizaje del Adaline, es aconsejable mostrar en pantalla el error medio a lo largo de los ciclos de aprendizaje (como se muestra en la siguiente tabla). El error de entrenamiento debe ir decreciendo o permanecer constante. El error de validación debe ir decreciendo, aunque pudiera aumentar, lo que significa que se está produciendo sobreaprendizaje.

Ciclo	Error medio de entrenamiento	Error medio de validación
1	2.1	2.8
2	1.5	1.9
3	0.9	1.7
.... etc

Tabla 1

Además de realizar el aprendizaje de la red, el programa debe:

- Calcular el error sobre el conjunto de test una vez finalizado el aprendizaje.
- Guardar en fichero las salidas de la red para todos los patrones de test. Estas salidas deberán “desnormalizarse”, es decir, pasarán del rango [0,1] al rango de valores originales. Para ello, serán necesarios los valores máximo y mínimo que se utilizaron para hacer la normalización de la salida.
- Guardar en fichero los pesos y el umbral de la red una vez finalizado el aprendizaje.
- Es conveniente también mostrar los errores de entrenamiento y validación al finalizar el aprendizaje, así como guardar en fichero las salidas para entrenamiento y validación por si fueran necesarios.

Con los datos procesados y el programa desarrollado, se realizarán **diferentes experimentos cambiando el valor de la razón de aprendizaje**, con el objetivo de encontrar el valor más adecuado para el problema dado. El valor óptimo se elegirá utilizando el error de validación. El número de ciclos de aprendizaje más adecuado para cada razón de aprendizaje hay que ajustarlo a cada experimento para conseguir la estabilización del error de entrenamiento y validación.

Resultados que deben mostrarse en la memoria de la práctica sobre el Adaline:

- *Tablas o gráficas de evolución de errores a lo largo del aprendizaje (sólo las más significativas de los experimentos realizados)*
- *Tabla resumen con los errores de entrenamiento, validación y test al acabar el aprendizaje para los experimentos realizados variando la razón de aprendizaje.*
- *Elegir el mejor experimento y realizar una gráfica que muestre la salida obtenida por la red y la salida deseada para los datos de test. Como se explicó anteriormente, estas salidas deberán “desnormalizarse” para que el rango de valores sea el original. Calcular también para este experimento el error en test con las salidas desnormalizadas para conocer la magnitud del error en MPa.*

3. Experimentación con el Perceptron Multicapa

Para el uso de Perceptron Multicapa se va a utilizar el **simulador snns** bajo el lenguaje de programación R. El paquete **RSNNS** permite un uso fácil de dicho simulador bajo el entorno de R. Se facilitará el script básico a utilizar para entrenar el Perceptron Multicapa, así como para calcular su error en diferentes conjuntos de datos o realizar un gráfico que muestre la evolución de los errores a lo largo de las iteraciones.

En la experimentación con el Perceptron Multicapa se realizarán diferentes pruebas, **cambiando el número de neuronas ocultas y la razón de aprendizaje**, con el objetivo de encontrar los valores más adecuados para el problema que se pretende resolver. Los valores óptimos se elegirán utilizando el error de validación. El número de experimentos a realizar siempre depende del problema, aunque se sugiere realizar al menos **tres experimentos cambiando el número de neuronas ocultas y otros tres cambiando la razón de aprendizaje**. El número de ciclos de aprendizaje más adecuado para cada configuración hay que ajustarlo a cada experimento para conseguir la estabilización del error de entrenamiento y validación.

Al igual que para el Adaline, durante el proceso de aprendizaje se utilizarán los conjuntos de entrenamiento y validación. Una vez finalizado el aprendizaje, se utilizará el conjunto de test para comprobar la capacidad de generalización de la red.

Resultados que deben mostrarse en la memoria de la práctica sobre el Perceptron Multicapa:

- *Tablas o gráficos de evolución de errores a lo largo del aprendizaje (sólo las más significativas de los experimentos realizados)*
- *Tabla resumen con los errores de entrenamiento, validación y test al acabar el aprendizaje para los experimentos realizados variando el número de neuronas ocultas y la razón de aprendizaje.*
- *Elegir el mejor experimento y realizar una gráfica que muestre las salidas desnormalizadas, tanto la obtenida por la red como la salida deseada para los datos de test. Calcular también para este experimento el error en test con las salidas desnormalizadas para conocer la magnitud del error en MPa.*

Entrega de la Práctica

Se entregará una memoria de la práctica en PDF, así como los ficheros de datos utilizados, los ficheros de resultados y los modelos entrenados.

La memoria deberá contener, al menos, un capítulo de introducción, un capítulo para cada uno de los modelos (Adaline y Perceptron Multicapa), con una descripción de toda la experimentación realizada, los resultados obtenidos y un análisis de los resultados, y un capítulo final con las conclusiones obtenidas al interpretar y comparar los resultados experimentales.

La entrega de la práctica se realizará la semana que se indica en el cronograma, sujeto a posibles variaciones que puedan producirse, en cuyo caso se avisará con suficiente antelación.

NOTA IMPORTANTE: Para poder comparar los resultados del Perceptron Multicapa con los resultados del Adaline, no hay que olvidar que los errores que se estén computando sean los mismos y no expresiones diferentes. Se recomienda utilizar el error medio absoluto (MAE) o el error cuadrático medio (MSE).