

OPENCOURSEWARE
REDES DE NEURONAS ARTIFICIALES
Javier Huertas Tato



Enunciado Caso Práctico V: SOM

1. Resumen

En esta práctica se aprenderá el uso de los Mapas autoorganizados de Kohonen (SOM) para tratar un problema de clasificación. Los SOM son habitualmente utilizados para resolver problemas no supervisados de clustering, pero en esta práctica se extenderá su uso a la clasificación.

En este enunciado se pide resolver un problema de clasificación usando el SOM. Se completará una experimentación haciendo uso del software SOM-PAK, ajustando los parámetros de dicho programa. Acto seguido se compararán con resultados de clasificación anteriores.

Al igual que en el resto de prácticas, también es necesario preprocesar los datos (revisar documento [PR-F-01]).

2. Materiales

Se necesitará:

- Dominio de datos de Jaen [PR-F-03.2] (Leer su descripción *Dominio 2 – Nubes Jaen*) [PR-F-03.1]
- Software SOM-PAK [PR-F-08.2]
- Guía del software SOM-PAK [PR-F-08.3]

3. Guía para la elaboración

Los pasos a seguir para completar la práctica son:

- Preprocesado de datos:
 - Normalizado de los datos entre 0 y 1
 - Aleatorización de la ordenación de los ejemplos de entrada
 - Separación de los datos en entrenamiento y test (80% y 20% respectivamente). Dado que este problema es de clasificación, es necesario que haya la misma proporción de clases en distintos conjuntos.
 - Respetar el formato de SOM PAK (Transparencia 1 de la Guía del software). Encabezado con el número de entradas. Separación por espacios. Decimales con puntos.
 - Los archivos que contengan los datos deben tener los siguientes nombres:
 - *DatosTrain.csv*
 - *DatosTest.csv*
- Uso de SOM-PAK:

- El software se ejecuta desde una consola de comandos (Windows o Linux).
- Los pasos a seguir se detallan a continuación:
 1. Inicializar aleatoriamente el mapa (*randinit*)
 2. Ajuste general del mapa (*vsom*)
 3. Ajuste fino del mapa (*vsom*)
 4. Calibrar el mapa (*vcal*)
- Con estos pasos el mapa está entrenado, a partir del mapa entrenado, se pueden evaluar sus métricas de error.
 - Clasificar los ejemplos de entrenamiento y de test (*visual*)
 - Evaluar el qerror del mapa. Solo mide lo buena que es la red agrupando elementos similares NO lo buena que es la red clasificando. (*qerror*)
 - Obtener visualizaciones de los mapas (*sammon* y *umat*)
- A partir de los ficheros clasificados se pueden calcular la precisión, precisión por clase y matriz de confusión.
- Experimentación con SOM
 - Los mapas autoorganizados tienen un conjunto de parámetros grande. Se pide probar los más influyentes.
 - Los requisitos mínimos de una experimentación son:
 - Probar 4 topologías, 3 bidimensionales y una unidimensional
 - Probar 3 ajustes de neuronas (con *vsom*) diferentes
 - Probar ambos tipos de topología (Hexagonal y rectangular).
 - Al final debería haber como mínimo 12 experimentos completos con la precisión y precisión por clase de resultado.
 - Cada experimento debe tener documentado su precisión y precisiones por clase.
 - El experimento final debe tener sus matrices de confusión, visualización gráfica y fichero de salidas y mapa guardados.
- Comparaciones
 - El problema de clasificación es comparable con las prácticas anteriores (PM y RBR, prácticas 3 y 4 respectivamente)
 - Se debe aportar una tabla comparativa.
- Extraer conclusiones
- Memoria
 - Realizar una memoria breve que muestre la elaboración de una experimentación completa y los pasos realizados durante la práctica.