

Introducción al Diseño de Experimentos para el Reconocimiento de Patrones

Capítulo 7: Herramientas

Curso de doctorado impartido por
Dr. Quiliano Isaac Moro
Dra. Aranzazu Simón Hurtado
Marzo 2004

Contenidos

1. Matlab
2. Aspirine
3. SNNS/JNNS
4. Weka
5. SOM / LVQ
6. GESL.
7. Consejos finales.

2

Matlab

- **Ventajas:**
 - Dispone de una amplia biblioteca de funciones y procedimientos, tanto de cálculo estadístico, como de regresión y clasificación.
 - La Toolbox de NN tiene un gran número de tipos de redes.
 - Programable.
 - Pueden crearse nuevos tipos de redes, o modificar los existentes.
 - Interfaz con lenguaje C.
 - Interfaz gráfico incluido.
- **Desventajas:**
 - Es un producto comercial.
 - ¡Cuidado con su uso sin licencia!

3

Aspirine/Migraines

- **Simulador de RNA, tipo MLP (incluye TDNN). Aprendizaje tipo BP.**
- **Ventajas:**
 - Genera el simulador en código C para compilar y ejecutar → ejecución rápida.
 - Incluye código fuente modificable.
 - Se pueden añadir nuevas funciones de transferencia, algoritmos de aprendizaje...
 - Tiene un interfaz de consola programable llamado Migraines.
 - Usa "pipes" tipo UNIX, lo que le da una gran versatilidad.
 - Se pueden construir sistemas modulares.
 - Incluye programas para: cálculo de Componentes Principales, cálculo de Análisis Canónico Factorial, evaluación de la clasificaciones.
 - Libre distribución.
- **Desventajas:**
 - Sólo MLP con BP.
 - La última versión conocida (V 6.0) es muy antigua (1992).
 - Versión 7 incluye un interfaz gráfico (2000).
 - Sólo entornos UNIX/LINUX.
 - Hubo intentos para el entorno MSWindows.

4

SNNS/JNNS

- **Simulador multiplataforma que incluye un modo consola y otro con interfaz gráfico.**
- **Ventajas:**
 - Es un software de libre distribución.
 - Gran número de: tipos de redes, funciones de activación, algoritmos de aprendizaje, funciones de inicialización de los pesos.
 - Definición de la red en modo gráfico / interactivo.
 - Visualización de: pesos y activaciones de los elementos de proceso, redes en 3D, curvas de error de aprendizaje, mapa de activación, ...
 - Puede ejecutarse en modo batch (consola), con su propio lenguaje de programación, aunque es un poco limitado.
 - Puede ejecutarse de manera distribuida.
 - Tiene herramientas para la evaluación de clasificadores.
 - Incluye técnicas de poda.
 - Después de haber entrenado la red se puede crear un fichero fuente C para ejecutar la red.
- **Desventajas:**
 - Es más lento que Aspirine/Migraines.
 - El acceso en modo batch no es tan eficiente como Aspirine/Migraines.
 - La última versión (4.2) fue realizada en el año 1998.
 - Últimamente hay una versión con el interfaz en JAVA (JNNS)

5

Weka

- **Paquete de Software escrito en JAVA para implementar algoritmos de aprendizaje.**
- **Ventajas:**
 - De libre distribución.
 - Multiplataforma.
 - Tiene muchos algoritmos de regresión/clasificación.
 - Incluye meta-algoritmos de aprendizaje (Bagging, AdaBoosting...)
 - Tiene preprocesado de datos (selección, estadísticas,...)
 - Incorpora herramientas para la visualización de los datos y resultados.
 - Se distribuye también su código fuente JAVA.
 - Se pueden añadir nuevas clases de clasificadores y filtros.
 - Tiene versiones de consola y con interfaz gráfico.
- **Desventajas:**
 - Lento.
 - Para RNA es muy limitado.

6

SOM/LVQ

- Paquetes de software muy pequeños, acordes con la simplicidad de los tipos de redes.
 - De libre distribución.
 - Incluye el código fuente.
 - Al ser programas cortos, son fáciles de modificar.
- Dispone de aplicaciones para visualizar los mapas creados.

7

GESL

8

Consejos finales

- En plataformas tipo UNIX el uso de Shell Scripts puede ser crucial para el diseño de experimentos complejos.
 - Herramientas como grep, awk, bc, ...
 - Creación de nuevas herramientas.
 - Uso intensivo del modo consola o batch y la redirección de entradas y salidas, así como las tuberías.
 - Problema: hay diferentes intérpretes shell, y diferentes versiones. NO está garantizada la portabilidad.
 - ¿Y si en vez de usar shell scripts se usa un lenguaje de programación como C?
- ¿Desarrollar la propia herramienta de clasificación?

9