

Sesión 1 - Tutorial. Entorno de trabajo

La primera sesión práctica plantea la correcta instalación y puesta en marcha de las herramientas software con las que trabajaremos en la asignatura: KNIME, WEKA y JADE. Para ello, se realiza una breve descripción de cada uno de las herramientas y se facilitan los recursos necesarios para acceder e instalar la versión adecuada para trabajar en función de las características del S.O. que se utilice.

1 KNIME

Una de las herramientas que vamos a utilizar en la asignatura es KNIME (*Konstanz Information Miner*). Se trata de una plataforma gratuita de minería de datos que permite construir modelos de aprendizaje utilizando un entorno visual desarrollado con Eclipse. La herramienta permite la confección de flujos de trabajo mediante nodos (que encapsulan distintos tipos de algoritmos) y flechas (que representan flujos de datos) que se despliegan y se pueden combinar interactivamente. Los nodos implementan distintos tipos de acciones que pueden ejecutarse sobre las tablas de datos:

- Manipulación de filas, columnas, muestreos, transformaciones, agrupaciones,...
- Visualización de datos y resultados (tablas dinámicas, histogramas, ...)
- Creación de modelos estadísticos y de minería de datos, como árboles de decisión, máquinas de vector soporte, regresiones, ...
- Validación de modelos, como curvas ROC, ...
- Evaluación y valoración de dichos modelos sobre conjuntos nuevos de datos.
- Creación de informes a medida.
- Creación de nodos específicos por parte del usuario mediante distintos lenguajes de programación.

La herramienta puede extenderse mediante la creación de nuevos nodos que implementen algoritmos y existe la posibilidad de usar directamente *WEKA* o incorporar código desarrollado en *R* o *python/jython*.

El sencillo interfaz de KNIME ha facilitado su uso a profesionales no relacionados con el ámbito de la informática y se utiliza en multitud de áreas como la investigación farmacéutica, análisis de datos de cliente de CRM, inteligencia de negocio y análisis de datos financieros.

Instalación

1. Comprobar versión del S.O. (Linux, Windows 32 o 64, Mac OS X)
2. Descargar ficheros del sitio WEB (<https://www.knime.com/downloads>). Hay varias posibilidades: Versión sencilla (sin extensiones), Versión completa (con extensiones) y Versión de desarrollo (incluye ECLIPSE).
3. Tenéis disponibles los siguientes tutoriales:
 - Tutorial básico: <https://youtu.be/ft7Ksgss3Tc>
 - Aprendiendo de los ejemplos: <https://youtu.be/qp1a0A40D6E>
 - Construir el primer Workflow: <https://youtu.be/F2cu2ewxxS0>

2 Weka

WEKA (*Waikato Environment for Knowledge Analysis*) es un entorno de software para aprendizaje automático y minería de datos escrito en Java y desarrollado en la Universidad de Waikato. Se trata de software libre distribuido bajo la licencia GNU-GPL. El software está completamente implementado en JAVA y contiene una colección de herramientas de visualización y algoritmos para análisis de datos y modelado predictivo.

Instalación

1. Instalar en KNIME la extensión de nodos Weka (ver tutorial).
2. Comprobar versión del S.O. (Linux, Windows 32 o 64, Mac OS X)
3. Descargar ficheros del sitio WEB (<http://www.cs.waikato.ac.nz/ml/weka/downloading.html>).
->Es necesario instalar antes en KNIME los nodos de tipo WEKA y luego descargar de la web la misma versión del entorno WEKA para que funcione desde KNIME correctamente.
4. Tenéis disponibles los siguientes tutoriales:
 - Instalación de extensiones en KNIME: <https://youtu.be/8HMx3mjJXiw>
 - Tutorial básico: <https://youtu.be/m7kpIBGEdkI>

3 JADE

JADE (*Java Agent DEvelopment Framework*) es un entorno software implementado completamente en JAVA para sistemas multiagente a través de un middleware que se ajusta a las especificaciones de FIPA (*Foundation for Physical Agents*). El entorno proporciona un conjunto de herramientas de tipo gráfico que facilitan el depurado y puesta en marcha de los sistemas.

Un sistema basado en JADE puede construirse de forma distribuida, utilizando distintas plataformas hardware. Su configuración puede controlarse a través de un GUI, incluso en tiempo de ejecución. Se distribuye con licencia LGPL (*Lesser General Public License Version 2*).

Instalación

La información relacionada con la instalación de la plataforma está referida en el siguiente documento que encontraréis descargar en Moodle:

SATD_Practicas_Tutorial_Instalacion_JADE.pdf