

TELA: environment for Ensemble Learning Analysis

Directores: Francisco Manuel Carrero y José Carlos Cortizo

Resumen: El objetivo de este proyecto es el de obtener una sistema de apoyo a la experimentación con conjuntos (ensembles) de clasificadores.

Contextualización: El aprendizaje de conjuntos de clasificadores es un área de gran interés para la comunidad investigadora en Aprendizaje Automático ya que muestran un rendimiento muy superior a los clasificadores simples en una gran variedad de entornos y dominios.

Sin embargo y debido a la relativa juventud de estas técnicas, no existe ninguna herramienta que permita una experimentación extensa, simple y rápida de estas técnicas en diversos dominios. Algunos acercamientos a un sistema similar incluyen a Weka que, en sus últimas versiones, incluye una opción para la generación de conjuntos de clasificadores pero muestra una interfaz muy poco intuitiva y escasas opciones de configuración de los conjuntos de clasificadores.

Objetivos: Los objetivos más importantes, a grandes rasgos, del sistema son los siguientes:

- El sistema debe ser modular, así como permitirá la adición de nuevos elementos de forma simple.
- El sistema debe ser capaz de combinar los distintos algoritmos mediante las técnicas de combinación básicas (voting, stacking, bagging, etc.) así como algunas nuevas que se propondrán al comienzo del proyecto.
- El sistema debe proporcionar algoritmos básicos de selección de atributos por clasificador.
- El sistema debe permitir cargar y guardar configuraciones de ensambles, así como su edición.
- El sistema debe ser capaz de tomar distintos tipos de fuentes de datos como origen de los mismos, así como integrarlos en una única.
- El sistema será capaz de ofrecer las siguientes salidas
 - Modelo del ensemble.
 - Las salidas típicas de Weka (accuracy, etc.)
 - Curvas ROC, así como AUC (área bajo la curva).
- El sistema debe proporcionar una interfaz gráfica de usuario usable, simple y que permita acceder a toda la funcionalidad del sistema de forma integrada. Esta interfaz será “drag and drop” y permitirá colocar los metaclasificadores, los clasificadores simples, las conexiones entre los mismos, las entradas de los datos, etc. de una forma simple y visual.