

## CLASIFICACIÓN AUTOMÁTICA DE DATOS IRREGULARES PARA DERIVACIÓN POSTOPERATORIA

Goddard J.C.; Martínez F.M.; Martínez A.E.; Cornejo J.M. - U.A.M.I.

Rufiner H.L.; Acevedo R.C. - U.N.E.R. (ARG.)

López I.L. - U.A.Ch.

**Objetivo:** encontrar alternativas para la clasificación automática con redes neuronales en los casos en que existen problemas con los datos de entrenamiento.

**Descripción:** en este trabajo se proponen una serie de opciones para los casos en que los métodos tradicionales de aprendizaje maquina resultan ineficaces para encarar la clasificación de datos reales. Tal es el caso de ejemplos con datos perdidos, distribución no uniforme de los patrones en las clases y grandes diferencias en los niveles de los atributos. Dichos tipos de datos son muy comunes en los problemas que se presentan el ámbito médico por diversas razones. Las redes neuronales han sido muy utilizadas para problemas de clasificación en esta área. Sin embargo en casos como el presentado los métodos clásicos de entrenamiento presentan dificultades debido a la naturaleza y características especialmente irregulares de los ejemplos. Especialmente el problema de la escasa cantidad de ejemplos para una de las clases es difícil de resolver y no se logra la convergencia a una solución satisfactoria. Existe la posibilidad de mejorar esta situación mediante distintas técnicas tales como escalar los patrones, 'rellenar' los valores perdidos y 'reforzar' los ejemplos de las clases minoritarias. Sin embargo dichas técnicas no resuelven por completo el problema.

Desde hace un tiempo se han presentado una serie de algoritmos basados en técnicas evolutivas o genéticas como alternativas a los métodos de búsqueda por gradiente descendente utilizados en problemas de optimización.

La tarea de clasificación elegida en este trabajo consiste en determinar a donde deben ser enviados los pacientes en recuperación postoperatoria. Esta base de datos es especialmente difícil de clasificar por los problemas que existen en los ejemplos.

**Metodología:** los datos fueron tomados de la *Escuela de Enfermería de la Universidad de Kansas* (USA) y han sido ampliamente utilizados en la bibliografía. Se han reportado una exactitud del orden de 50 % con métodos de aprendizaje automático sin preproceso de los datos. Se cuenta con 90 ejemplos de 8 atributos y tres posibles clases. Dado que la hipotermia es un factor significativo luego de una cirugía los atributos corresponden en su mayoría a medidas derivadas de la temperatura corporal. Uno de los atributos posee 3 valores perdidos y una de las clases posee menos del 2% de los ejemplos. Asimismo la mayoría de los atributos consisten en valores discretos entre 0 y 1, mientras que uno de los atributos consiste en un valores discretos entre 0 y 20.

**Resultados:** se evaluaron distintas alternativas de algoritmos evolutivos combinados con técnicas de búsqueda por gradiente. En cada caso se analizó el desempeño en base a los errores cometidos en la clasificación y la velocidad de aprendizaje de los algoritmos.