

Árboles de Decisión

García Prado, Sergio

2 de marzo de 2017

Resumen

Abstract.

1. ¿POR QUÉ NO SE PUEDE APLICAR DIRECTAMENTE ID3?
2. REALICE LAS MODIFICACIONES PREVIAS EN EL FICHERO DE DATOS PARA QUE PUEDA LLEVAR A CABO LO ANTERIOR Y PROPORCIONE LOS RESULTADOS APLICANDO EL MÉTODO DE RETENCIÓN O HOLD-OUT PARA LA FORMACIÓN DEL EXPERIMENTO
3. VOLVIENDO AL FICHERO ORIGINAL, PASE EL ALGORITMO J48 TAMBIÉN APLICANDO EL MÉTODO DE RETENCIÓN O HOLD-OUT. ANALICE EL ÁRBOL OBTENIDO SIN PODA. ¿SE PODRÍA PRESCINDIR DE ALGÚN ATRIBUTO? SI ES ASÍ, HÁGALO Y COMPARE LOS RESULTADOS DE NUEVO
4. HABRÁ NOTADO QUE CUANDO SE USA ALGÚN ATRIBUTO NUMÉRICO, IMPLÍCITAMENTE SE APLICA UNA DISCRETIZACIÓN AL PLANTEAR LAS DIFERENTES RAMAS DEL ÁRBOL A PARTIR DE ÉL. ¿POR QUÉ ES MÁS EFICIENTE ESTA TÉCNICA QUE LA APLICADA EN 2?
5. PLANTEE, ENTONCES, UNA DISCRETIZACIÓN BASADA EN EL PUNTO ANTERIOR, AUNQUE NO RESULTE SER BINARIA, SINO EN TANTOS TRAMOS COMO INDUZCAN LOS VALORES USADOS AL FORMAR LAS RAMAS DEL ÁRBOL CON J48 SIN PODA (HOLD-OUT). INTRODUZCA ESTE FICHERO DE NUEVO AL ALGORITMO ID3. COMPARE LOS RESULTADOS CON EL J48 DE 3

REFERENCIAS

- [1] CALONGE CANO, T., AND ALONSO GONZÁLEZ, C. J. Técnicas de Aprendizaje Automático, 2016/17.