

Comparación entre J48(C4.5) y Naive Bayes

Sergio García Prado

7 de noviembre de 2016

Resumen

I. INTRODUCCIÓN

II. TEST DE McNEMAR: HOLDOUT DE 2/3

h_A es J48 y h_B es Naive Bayes

| | |
|--|--|
| Número de ejemplos mal clasificados por h_A y h_B (n_{00}) | Número de ejemplos mal clasificados por h_A pero no por h_B (n_{01}) |
| Número de ejemplos mal clasificados por h_B pero no por h_A (n_{10}) | Número de ejemplos bien clasificados por h_A y h_B (n_{11}) |

$$\frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}}$$

I. Soybean

| | |
|----|-----|
| 9 | 9 |
| 11 | 204 |

II. Vote

| | |
|----|-----|
| 4 | 4 |
| 13 | 127 |

III. Labor

Puesto que este conjunto de estadísticos es demasiado pequeño

| | |
|---|----|
| 2 | 2 |
| 0 | 16 |

III. TEST DE STUDENT: CROSS VALIDATION SIN REPETICIÓN

IV. TEST DE STUDENT: CROSS VALIDATION CON REPETICIÓN

V. RESULTADOS

Los resultados obtenidos según los test realizados con los conjuntos de datos y los tipos de test son los siguientes:

| | McNemar | Student | Student (C) | Student rep. | Student rep.(C) |
|---------|---------|---------|-------------|--------------|-----------------|
| Soybean | J48 | NB | NB | NB | NB |
| Vote | J48 | J48 | J48 | J48 | J48 |
| Labor | NB | NB | NB | NB | NB |