

Evaluación de Hipótesis

García Prado, Sergio
sergio@garciparedes.me

19 de marzo de 2017

Resumen

[TODO]

1. INTRODUCCIÓN

[TODO]

1.1. ALGORITMOS

[TODO]

- **J48**:
- **JRIP**:

1.2. CONJUNTOS DE DATOS

[TODO]

- **Labor**[data]:
- **Soybean**[datb]:
- **Vote**[datc]:

2. REALIZAR UN EXPERIMENTO APLICANDO HOLDOUT $\frac{2}{3}/\frac{1}{3}$

[TODO]

| Holdout 2/3, 1/3 | | |
|------------------|-----------|----------------------|
| Datos | Algoritmo | Tasa de Error |
| | | Semilla ₁ |
| Labor | J48 | 0,105263 |
| | JRIP | 0,105263 |
| Soybean | J48 | 0,094828 |
| | JRIP | 0,086207 |
| Vote | J48 | 0,027027 |
| | JRIP | 0,033784 |

Tabla 1

3. REALIZAR TRES EXPERIMENTOS ADICIONALES APLICANDO HOLDOUT $\frac{2}{3}/\frac{1}{3}$, ANOTANDO LA TASA DE ERROR DE CADA EXPERIMENTO

[TODO]

| Holdout 2/3, 1/3 Repetido | | | | |
|---------------------------|-----------|----------------------|----------------------|----------------------|
| Datos | Algoritmo | Tasa de Error | | |
| | | Semilla ₂ | Semilla ₃ | Semilla ₄ |
| Labor | J48 | 0,157895 | 0,315789 | 0,105263 |
| | JRIP | 0,157895 | 0,210526 | 0,105263 |
| Soybean | J48 | 0,112069 | 0,107759 | 0,137931 |
| | JRIP | 0,077586 | 0,116379 | 0,073276 |
| Vote | J48 | 0,081081 | 0,054054 | 0,060811 |
| | JRIP | 0,054054 | 0,047297 | 0,047297 |

Tabla 2

4. SOBRE LOS RESULTADOS CALCULADOS EN LA SECCIÓN 3 DETERMINARLA TASA DE ERROR, LA VARIANZA Y EL INTERVALO DE CONFIANZA DEL 95 %

[TODO]

| Holdout 2/3, 1/3: Global | | | | |
|--------------------------|-----------|---------------|---------------------|----------------------|
| Datos | Algoritmo | Tasa de Error | Desviación Estandar | Intervalos |
| Labor | J48 | 0,192982 | 0,109561 | [0,008277, 0,377686] |
| | JRIP | 0,157894 | 0,052631 | [0,069165, 0,246622] |
| Soybean | J48 | 0,119253 | 0,016318 | [0,091743, 0,146762] |
| | JRIP | 0,089080 | 0,023739 | [0,049059, 0,129100] |
| Vote | J48 | 0,065315 | 0,014065 | [0,041603, 0,089026] |
| | JRIP | 0,049549 | 0,003901 | [0,042972, 0,056125] |

Tabla 3

5. REALIZAR UN EXPERIMENTO DE VALIDACIÓN CRUZADA DE 10 PARTICIONES, CALCULANDO LA TASA DE ERROR

[TODO]

| Validación Cruzada | | |
|--------------------|-----------|----------------------|
| Datos | Algoritmo | Tasa de Error |
| | | Semilla ₁ |
| Labor | J48 | 0,263158 |
| | JRIP | 0,228070 |
| Soybean | J48 | 0,084919 |
| | JRIP | 0,077599 |
| Vote | J48 | 0,036782 |
| | JRIP | 0,045977 |

Tabla 4

6. REALIZAR TRES EXPERIMENTOS DE VALIDACIÓN CRUZADA DE 10 PARTICIONES, ANOTANDO LA TASA DE ERROR

[TODO]

| Validación Cruzada Repetida | | | | |
|-----------------------------|-----------|----------------------|----------------------|----------------------|
| Datos | Algoritmo | Tasa de Error | | |
| | | Semilla ₂ | Semilla ₃ | Semilla ₄ |
| Labor | J48 | 0,263158 | 0,263158 | 0,245614 |
| | JRIP | 0,140351 | 0,157895 | 0,157895 |
| Soybean | J48 | 0,098097 | 0,090776 | 0,079063 |
| | JRIP | 0,086384 | 0,068814 | 0,081991 |
| Vote | J48 | 0,032184 | 0,036782 | 0,034483 |
| | JRIP | 0,043678 | 0,041379 | 0,03908 |

Tabla 5

7. SOBRE LOS RESULTADOS CALCULADOS EN LA SECCIÓN 6 DETERMINARLA TASA DE ERROR

[TODO]

| Validación Cruzada Repetida: Global | | |
|-------------------------------------|-----------|---------------|
| Datos | Algoritmo | Tasa de Error |
| Labor | J48 | 0,25731 |
| | JRIP | 0,152047 |
| Soybean | J48 | 0,089312 |
| | JRIP | 0,079063 |
| Vote | J48 | 0,034483 |
| | JRIP | 0,041379 |

Tabla 6

8. CONCLUSIONES

[TODO]

| Conjunto de Datos: Labor | | | | |
|--------------------------|----------|------------------|--------------------|-----------------------------|
| Algoritmo | Holdout | Holdout Repetido | Validación Cruzada | Validación Cruzada Repetida |
| J48 | 0,105263 | 0,192982 | 0,263158 | 0,25731 |
| JRIP | 0,105263 | 0,157894 | 0,228070 | 0,152047 |

Tabla 7

| Conjunto de Datos: Soybean | | | | |
|----------------------------|----------|------------------|--------------------|-----------------------------|
| Algoritmo | Holdout | Holdout Repetido | Validación Cruzada | Validación Cruzada Repetida |
| J48 | 0,094828 | 0,119253 | 0,084919 | 0,089312 |
| JRIP | 0,086207 | 0,089080 | 0,077599 | 0,079063 |

Tabla 8

| Conjunto de Datos: Vote | | | | |
|-------------------------|----------|------------------|--------------------|-----------------------------|
| Algoritmo | Holdout | Holdout Repetido | Validación Cruzada | Validación Cruzada Repetida |
| J48 | 0,027027 | 0,065315 | 0,036782 | 0,034483 |
| JRIP | 0,033784 | 0,049549 | 0,045977 | 0,041379 |

Tabla 9

REFERENCIAS

- [CCAG17] Teodoro Calonge Cano and Carlos Javier Alonso González. Técnicas de Aprendizaje Automático, 2016/17.
- [data] Labor Data Set. <http://storm.cis.fordham.edu/~gweiss/data-mining/weka-data/labor.arff>.
- [datb] Soybean Data Set. <http://storm.cis.fordham.edu/~gweiss/data-mining/weka-data/soybean.arff>.
- [datc] Vote Data Set. <http://storm.cis.fordham.edu/~gweiss/data-mining/weka-data/vote.arff>.
- [GP17] Sergio García Prado. Técnicas de aprendizaje automático: Evaluación de Hipótesis. <https://github.com/garciparedes/machine-learning-hypothesis-evaluation>, 2017.
- [too] Weka. <http://www.cs.waikato.ac.nz/ml/weka/>.