



Taller:

# Clasificación con RBs

## MATLABArsenal y BNT

*Universidad de Castilla-La Mancha*  
03/02/2010

Ana M. Martínez

anamaria.martinez@uclm.es

Departamento de Sistemas Informáticos  
Albacete - UCLM - Spain



### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

### Weka (Java)

### Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

### Ejercicios Adicionales

### Referencias

## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

## ¿Por qué MATLABArsenal?

- Paquete en Matlab que encapsula diversos algoritmos de clasificación [ars].
  - Autor: Rong Yan, Carnegie Mellon University en Pensilvania.
- ▲ **Código fuente disponible.**
- ▲ Proporciona un entorno para la **evaluación** de dichos algoritmos: interfaces de entrada y salida estandarizadas para poder comparar el rendimiento entre distintos algoritmos de aprendizaje.
- ▲ Proporciona **interfaces** para el uso de otras herramientas de Minería de datos como **Weka**, **SVM\_light**, **libSVM** y **Netlab**.
- ▲ Preparado para manejar datos sparse.
- ▲ Extensible a incorporar nuevos algoritmos.
- ▼ Versión 1.0 Beta de marzo de 2006.
- ▼ Sólo datos numéricos :(.

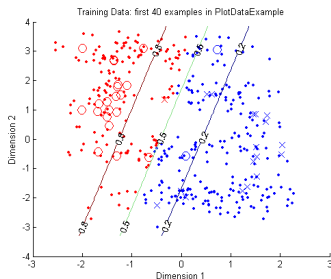


- 1 Algoritmos de Clasificación:** regresión logística (kernel), modelo de máxima entropía (IIS), análisis discriminante lineal (kernel), modelos de mixturas Gaussianas, Procesos Gaussianos, kNN, (Kernel) perceptrón, decision Stump.
- 2 Interfaz a otras herramientas de minería de datos:** Weka, SVM\_light, libSVM y Netlab.
- 3 Multiclasificadores:** AdaBoost, Bagging, Muestreo hacia adelante y hacia atrás, clasificación jerárquica, por voto, multiclase y multietiqueta.
- 4 Métodos de evaluación:** división en conjunto de entrenamiento y test, validación cruzada, o sólo conjunto de entrenamiento o predicción.



## Características que ofrece MATLABArsenal II

- 5 **Aprendizaje activo** para clases binarias y múltiples con estrategias del mejor y/o el peor caso.
- 6 Representación **sparse** de los datos.
- 7 **Preprocesado de bases de datos:** selección de atributos usando SVD, FLD o Latent Dirichlet Allocation (esta última para atributos binarios solamente).
- 8 **Visualización de las fronteras de decisión.**



## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias





## • Instalación en Windows:

- 1 Descargar el paquete desde <http://www.informedia.cs.cmu.edu/yanrong/MATLABArsenal/MATLABArsenal.zip>.
- 2 Descomprimir el .zip en un directorio, por ejemplo \$MATLABArsenalRoot.
- 3 Añadir la ruta \$MATLABArsenalRoot y sus subdirectorios en MATLAB (utilizar el comando `addpath` o en el menú `File->Set Path`).

## • Instalación en Linux:

- Todos los pasos anteriores y además: (siguiente transparencia)

MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

## ...Instalación en Linux

### 1 Interfaz WEKA:

- 1 Descomprimir `weka.jar` en el directorio `weka-3-4`.
- 2 Editar el fichero `Test_Classify.m`:  

```
%weka_dir = sprintf('%s/weka-3-4/weka.jar', root);
```

por:  

```
weka_dir = sprintf('%s/Weka-3-4/weka', root);
```

### 2 Interfaz libSVM:

- 1 <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>  
Ir a Download LIBSVM y descargar el archivo tar.gz.
- 2 En el directorio donde se descomprima ejecutar `make` en el bash.
- 3 Copiar los archivos: `svm-train`, `svm-predict` y `svm-scale` en la carpeta `svm` y renombrarlos quitándoles el guión.

### 3 Interfaz SVM\_light:

- 1 <http://svmlight.joachims.org/>  
En Source Code and Binaries descargar los binarios para Linux:  

```
http://download.joachims.org/svm_light/current/svm_light_linux.tar.gz
```
- 2 Copiar los archivos `svm_learn` y `svm_classify` en la carpeta `svm` de Arsenal.



## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias



- El módulo principal de MATLABArsenal es “Arsenal” (que redirige a Test\_Classify).
- ```
Arsenal('classify -t input_file [options]
        [--@Evaluation [options]] ...
        -- Classifier [param] [-- Classifiers]');
```
- Se indica el archivo de entrada (con sus opciones), el modo de evaluación, el clasificador utilizado (puede ser un multclasificador u otra herramienta).

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias



**Tabla:** Opciones sobre los tres **formatos de entrada**.

|                  |                                                                                                                                                                                                                    |
|------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>-if 0/1/2</b> | Utilizar el primer tipo, el segundo o el tercero:<br>-0 <valor1> <valor2> ... <valorn> <clase><br>-1 <fila> <columna> <valor><br>-2 <clase> <atributo1>:<valor1> <atributo2>:<br><valor2> ... <atributon>:<valorn> |
|------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

- **Ejemplo:** Primer tipo (-if 0)

```
0.3  4  3.5  0
0.4  4  3.8  0
0.7  5  4.2  1
```

- Dos **archivos de salida**, por defecto:

- 1 `$(input_file).pred`: contiene los resultados de las predicciones para cada instancia de test:  
`<instancia><prob.><pred.><real>`
- 2 `$(input_file).result`: contiene los estadísticos de predicción para el conjunto de test: resumen de las opciones de clasificación, error, precisión, recall (exactitud), precisión media (MAP) y otras medidas.

\* Nota:

- $Tasa\ de\ acierto(accuracy) / error = \frac{tp+tn}{tp+tn+fp+fn}$
- $Precision = \frac{tp}{tp+fp}$
- $Recall(exactitud) = \frac{tp}{tp+fn}$
- $F1 = \frac{2*prec*rec}{prec+rec}$
- MAP: precisión media.
- MBAP: Línea base aleatoria con respecto a la precisión media.

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

- `Arsenal('classify -t input_file [options]  
          [--@Evaluation [options]] ...  
          -- Classifier [param] [-- Classifiers]');`



**Tabla:** Opciones generales (preprocesado).

|        |           |                                                         |
|--------|-----------|---------------------------------------------------------|
| -v     | (def 1)   | Mostrar todos los mensajes (1) o no (0)                 |
| -sf    | (def 0)   | Barajar los datos (1) o no (0).                         |
| -n     | (def 1)   | Normalizar los datos (1) o no.                          |
| -svd   | (def 0)   | reducción de la dimensión SVD (grado reducción).        |
| -fld   | (def 0)   | reducción de la dimensión Fisher LDA (grado reducción). |
| -ldir  | (def 0)   | reducción de la dimensión LDA (rango: {0, 1}).          |
| -map   | (def 0)   | mostrar la precisión media (MAP).                       |
| -if    | (def 0)   | formato de entrada.                                     |
| -of    | (def 0)   | formato de salida.                                      |
| -pf    | (def 0)   | formato de la predicción: 0 ó 1.                        |
| -chi   | (def 0)   | selección de atributos ChiSquare.                       |
| -t     | (def '')  | nombre del archivo de entrada.                          |
| -o     | (def '')  | nombre del archivo de salida.                           |
| -p     | (def '')  | nombre del archivo de predicción.                       |
| -oflag | (def 'a') | flag de salida: 'a' agregar,'w' sobrecribir.            |
| -dir   | (def '')  | directorio de trabajo, que es \$MATLABArsenalRoot       |
| -drf   | (def '')  | archivo intermedio para reducción de la dimensión.      |
| -label | (def if)  | conjunto de clases a procesar.                          |

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias



Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

- Arsenal('classify -t input\_file [options]  
[-@Evaluation [options]] ...  
-- Classifier [param] [-- Classifiers]');



**Tabla:** Opciones métodos de evaluación.

|                              |                                                                                                                      |
|------------------------------|----------------------------------------------------------------------------------------------------------------------|
| train_test_validate<br>(def) | división datos de entrada en entrenamiento y test.<br>opciones: -t (def -2, la mitad),<br>punto corte entrenamiento. |
| cross_validate               | Validación cruzada.<br>opciones: -t (def -3), n <sup>o</sup> particiones.                                            |
| test_file_validate           | entrenamiento y test por separado<br>opciones: -t (def ""), archivo de test adicional                                |
| train_only                   | sólo entrenamiento<br>opciones: -m (def ""), modelo de salida                                                        |
| test_only                    | sólo test<br>opciones: -m (def ""), modelo de entrada                                                                |

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

## Ejercicios Adicionales

## Referencias

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje

estructural de Leray y

François

Ejemplos

Ejercicios Adicionales

Referencias

- `Arsenal('classify -t input_file [options]`  
    `[--@Evaluation [options]] ...`  
    `-- Classifier [param] [-- Classifiers]');`



**Tabla:** Opciones algoritmos multiclase.

|                                           |                                                                                                                                                                                                                                                                                                                                                                                                               |
|-------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>train_test_multiple_class</code>    | Clasificación multiclase<br>opciones: <code>-CodeType</code> (def 0), 0: uno contra todos, 1: Emparejamiento por pares, 2: ECOC-16<br><code>-LossFuncType</code> (def 2), 0: función de pérdida logística, 1: exponencial, 2: hinge loss.                                                                                                                                                                     |
| <code>train_test_multiple_class_AL</code> | Clasificación multiclase con aprendizaje activo<br>opciones: <code>-CodeType</code> (def 0), 0: uno contra todos, 1: Emparejamiento por pares, 2: ECOC-16<br><code>-LossFuncType</code> (def 2), 0: función de pérdida logística, 1: exponencial, 2: hinge loss.<br><code>-ALIter</code> (def 4), iteraciones para el aprend. activo.<br><code>-ALIncrSize</code> (def 10): Tamaño incremental por iteración. |

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## Ejercicios Adicionales

## Referencias

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje

estructural de Leray y

François

Ejemplos

Ejercicios Adicionales

Referencias

- `Arsenal('classify -t input_file [options]`  
    `[--@Evaluation [options]] ...`  
    `-- Classifier [param] [-- Classifiers']');`



**Tabla:** Opciones multclasificadores.

|                     |                                                                                                                                                                                                                          |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| WekaClassify        | (para) - - (+ clasif. y opciones)<br>-MultiClassWrapper (def -1): 1 activado,<br>0 desactivado, -1 selección automática.                                                                                                 |
| MCActiveLearning    | Módulo de aprendizaje activo.<br>-Iter (def 10): n <sup>o</sup> iteraciones.<br>-IncrSize (def 10): tamaño incremental/iteración.                                                                                        |
| MCAdaBoostM1        | AdaBoost.M1<br>-Iter (def 10): n <sup>o</sup> iteraciones.<br>-SampleRatio (def 1): 1 significa 100 % de los datos.                                                                                                      |
| MCSBagging          | Bagging<br>-Iter (def 10): n <sup>o</sup> iteraciones.<br>-SampleRatio (def 1): 1 significa 100 % de los datos.                                                                                                          |
| MCDownSampling      | Muestreo hacia atrás.<br>-PosRatio (def 0.5): ratio de datos positivos al acabar.                                                                                                                                        |
| MCUpSampling        | Muestreo hacia adelante.<br>-PosRatio (def 0.5): ratio de datos positivos al acabar.                                                                                                                                     |
| MCHierarchyClassify | (para) - - (Meta Clasificador, para) [- - ClasificadorBase]<br>-PosRatio (def 0.5): ratio de datos positivos al acabar.<br>-SampleDevSet (def 0): usar un conjunto muestral para aprender el meta clasificador (1) o no. |
| MCWithMultiFSet     | Clasificación jerárquica en varios grupos de atributos<br>-Voting (def 0): regla de la suma o la mayoría, 0 es suma.<br>-Separator (def 0): separadores para grupos de atributos                                         |

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## Ejercicios Adicionales

## Referencias



**Tabla:** Opciones clasificadores base I.

|                              |                                                                                                                                                                                                                                                                                                 |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SVM_LIGHT                    | -Kernel (def 0): 0 lineal, 1 polinomial, 2 RBF y 3 sigmoidal.<br>-KernelParam (def 0.05).<br>-CostFactor (def 1): ratio de datos positivos y negativos.<br>-Threshold (def 0): clasificado como + si mayor que el umbral.<br>-TransPosFrac (def -1): Si $> 0$ aprendizaje transductivo trigger. |
| libSVM                       | -Kernel (def 0): 0 lineal, 1 polinomial, 2 RBF y 3 sigmoidal.<br>-KernelParam (def 0.05).<br>-CostFactor (def 1): ratio de datos positivos y negativos.<br>-Threshold (def 0): clasificado como + si mayor que el umbral.                                                                       |
| kNN_classify                 | -k (def 1): número de vecinos.<br>-d (def 2): 0 dist. Euclidean, 1 chi-squared y 2 similaridad cos.                                                                                                                                                                                             |
| GMM_classify                 | -NumMix (def 1): n° de mixturas para cada clase.                                                                                                                                                                                                                                                |
| LDA_classify                 | -RegFactor (def 0.1).<br>-QDA (def 0): 0 para LDA, 1 para QDA.                                                                                                                                                                                                                                  |
| IIS_classify                 | -lter (def 50): n° de iteraciones<br>-MinDiff (def $1e - 7$ ): diferencia mín. de la logverosimilitud.<br>-Sigma (def 0): factor de regulación.                                                                                                                                                 |
| NeuralNet<br>(Perc. Multic.) | -NHidden (def 10): unidades ocultas.<br>-NOut (def 1): unidades de salida.<br>-Alpha (def 0.2): peso<br>-NCycles (def 10): número de ciclos de entrenamiento.<br>-Threshold (def 0.5): umbral de clasificación.                                                                                 |

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje

estructural de Leray y

François

Ejemplos

## Ejercicios Adicionales

## Referencias



**Tabla:** Opciones clasificadores base II.

|                  |                                                                                                                                                                                                                                  |
|------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| GP_Classify      | -PriorMean (def 0): media de la función a priori.<br>-PriorVariance (def 1): varianza de la función a priori.<br>-NCycles (def 10): n <sup>o</sup> de ciclos de entrenamiento.<br>-Threshold (def 0.5): umbral de clasificación. |
| LogitReg         | -RegFactor (def 0).<br>-CostFactor (def 1).                                                                                                                                                                                      |
| LogitRegKernel   | -RegFactor (def 0).<br>-Kernel (def 0): 0 lineal, 1 polinomial, 2 RBF y 3 sigmoideal.<br>-KernelParam (def 0.05).                                                                                                                |
| Perceptron       | -MaxIter (def 100): n <sup>o</sup> máximo de iteraciones.<br>-CostFactor (def 1).                                                                                                                                                |
| PerceptronKernel | -MaxIter (def 100): n <sup>o</sup> máximo de iteraciones.<br>-CostFactor (def 1).<br>-Kernel (def 0): 0 lineal, 1 polinomial, 2 RBF y 3 sigmoideal.<br>-KernelParam (def 0.05).                                                  |
| DecisionStump    | -Threshold (def 0): umbral de clasificación.<br>-CostFactor (def 1).                                                                                                                                                             |
| ZeroR            | todo se predice como 0.                                                                                                                                                                                                          |

### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias



## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

# Ejemplos (Clasificadores basados en Redes Bayesianas)

- Clasificador Naive Bayes: validación cruzada 10 particiones.

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje

estructural de Leray y

François

Ejemplos

Ejercicios Adicionales

Referencias

- Clasificador Naive Bayes: validación cruzada 10 particiones.

```
Arsenal('classify -t iris.data -sf 1 -- cross_validate -t 10  
-- WekaClassify -MultiClassWrapper 0 -- bayes.NaiveBayes');
```

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias



- Clasificador Naive Bayes: validación cruzada 10 particiones.

```
Arsenal('classify -t iris.data -sf 1 -- cross_validate -t 10  
-- WekaClassify -MultiClassWrapper 0 -- bayes.NaiveBayes');
```

```
Finished loading iris.data.....  
Output Results to iris.data.result.....  
Output Predictions to iris.data.pred.....  
Message: Cross Validation, Folder: 10, Classification, Shuffled  
Data Number: 150, Feature Number: 4, Data per Class: (1,50) (2,50) (3,50)  
Iteration 1 .....  
Classifier: NaiveBayes, Classifier: bayes.NaiveBayes, Parameters: -MultiClassWrapper: 0  
Prec:1.000000, Rec:1.000000, Err:0.000000  
Iteration 2 .....  
  
Iteration 10 .....  
Classifier: NaiveBayes, Classifier: bayes.NaiveBayes, Parameters: -MultiClassWrapper: 0  
Prec:0.800000, Rec:0.800000, Err:0.200000  
Error = 0.053333, Precision = 0.946667, Recall = 0.946667, F1 = 0.946667, MAP = 0.000000,  
MBAP = 0.000000, Time = 3.205766,
```

- Ó de forma equivalente:

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

## Ejercicios Adicionales

## Referencias



- Clasificador Naive Bayes: validación cruzada 10 particiones.

```
Arsenal('classify -t iris.data -sf 1 -- cross_validate -t 10  
-- WekaClassify -MultiClassWrapper 0 -- bayes.NaiveBayes');
```

```
Finished loading iris.data.....  
Output Results to iris.data.result.....  
Output Predictions to iris.data.pred.....  
Message: Cross Validation, Folder: 10, Classification, Shuffled  
Data Number: 150, Feature Number: 4, Data per Class: (1,50) (2,50) (3,50)  
Iteration 1 .....  
Classifier: NaiveBayes, Classifier: bayes.NaiveBayes, Parameters: -MultiClassWrapper: 0  
Prec:1.000000, Rec:1.000000, Err:0.000000  
Iteration 2 .....  
  
Iteration 10 .....  
Classifier: NaiveBayes, Classifier: bayes.NaiveBayes, Parameters: -MultiClassWrapper: 0  
Prec:0.800000, Rec:0.800000, Err:0.200000  
Error = 0.053333, Precision = 0.946667, Recall = 0.946667, F1 = 0.946667, MAP = 0.000000,  
MBAP = 0.000000, Time = 3.205766,
```

- Ó de forma equivalente:

```
Arsenal('classify -t iris.data -sf 1 -- cross_validate -t 10  
--NaiveBayes');
```

- Conjunto de entrenamiento para validar el modelo:

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

## Ejercicios Adicionales

## Referencias



- Clasificador Naive Bayes: validación cruzada 10 particiones.

```
Arsenal('classify -t iris.data -sf 1 -- cross_validate -t 10  
-- WekaClassify -MultiClassWrapper 0 -- bayes.NaiveBayes');
```

```
Finished loading iris.data.....  
Output Results to iris.data.result.....  
Output Predictions to iris.data.pred.....  
Message: Cross Validation, Folder: 10, Classification, Shuffled  
Data Number: 150, Feature Number: 4, Data per Class: (1,50) (2,50) (3,50)  
Iteration 1 .....  
Classifier: NaiveBayes, Classifier: bayes.NaiveBayes, Parameters: -MultiClassWrapper: 0  
Prec:1.000000, Rec:1.000000, Err:0.000000  
Iteration 2 .....  
  
Iteration 10 .....  
Classifier: NaiveBayes, Classifier: bayes.NaiveBayes, Parameters: -MultiClassWrapper: 0  
Prec:0.800000, Rec:0.800000, Err:0.200000  
Error = 0.053333, Precision = 0.946667, Recall = 0.946667, F1 = 0.946667, MAP = 0.000000,  
MBAP = 0.000000, Time = 3.205766,
```

- Ó de forma equivalente:

```
Arsenal('classify -t iris.data -sf 1 -- cross_validate -t 10  
--NaiveBayes');
```

- Conjunto de entrenamiento para validar el modelo:

```
Arsenal('classify -t weather.data -- Test_File_Validate -t weather.data --NaiveBayes');
```

- Ó de forma equivalente:

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

## Ejercicios Adicionales

## Referencias



- Clasificador Naive Bayes: validación cruzada 10 particiones.

```
Arsenal('classify -t iris.data -sf 1 -- cross_validate -t 10  
-- WekaClassify -MultiClassWrapper 0 -- bayes.NaiveBayes');
```

```
Finished loading iris.data.....  
Output Results to iris.data.result.....  
Output Predictions to iris.data.pred.....  
Message: Cross Validation, Folder: 10, Classification, Shuffled  
Data Number: 150, Feature Number: 4, Data per Class: (1,50) (2,50) (3,50)  
Iteration 1 .....  
Classifier: NaiveBayes, Classifier: bayes.NaiveBayes, Parameters: -MultiClassWrapper: 0  
Prec:1.000000, Rec:1.000000, Err:0.000000  
Iteration 2 .....  
  
Iteration 10 .....  
Classifier: NaiveBayes, Classifier: bayes.NaiveBayes, Parameters: -MultiClassWrapper: 0  
Prec:0.800000, Rec:0.800000, Err:0.200000  
Error = 0.053333, Precision = 0.946667, Recall = 0.946667, F1 = 0.946667, MAP = 0.000000,  
MBAP = 0.000000, Time = 3.205766,
```

- Ó de forma equivalente:

```
Arsenal('classify -t iris.data -sf 1 -- cross_validate -t 10  
--NaiveBayes');
```

- Conjunto de entrenamiento para validar el modelo:

```
Arsenal('classify -t weather.data -- Test_File_Validate -t weather.data --NaiveBayes');
```

- Ó de forma equivalente:

```
Arsenal('classify -t weather.data -- Train_Only -m weather.model --NaiveBayes');  
Arsenal('classify -t weather.data -- test_only -m weather.model --NaiveBayes');
```

(Mismo resultado que en Weka con Gaussianas)

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

## Ejercicios Adicionales

## Referencias

## Ejemplos (Otros algoritmos de minería de datos)

- Barajar datos de entrada, reducción  $n^{\circ}$  de atributos a 15, validación cruzada 3 particiones con el clasificador 3NN.



### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

### Weka (Java)

### Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

### Ejercicios Adicionales

### Referencias



## Ejemplos (Otros algoritmos de minería de datos)

- Barajar datos de entrada, reducción  $n^{\circ}$  de atributos a 15, validación cruzada 3 particiones con el clasificador 3NN.

```
Arsenal('classify -t DataExample1.txt -sf 1 -svd 15 -drf  
DataExample1svd.txt --cross_validate -t 3 --kNN_classify -k 3');
```



### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

### Weka (Java)

### Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

### Ejercicios Adicionales

### Referencias

## Ejemplos (Otros algoritmos de minería de datos)

- Barajar datos de entrada, reducción  $n^{\circ}$  de atributos a 15, validación cruzada 3 particiones con el clasificador 3NN.

```
Arsenal('classify -t DataExample1.txt -sf 1 -svd 15 -drf  
DataExample1svd.txt --cross_validate -t 3 --kNN_classify -k 3');
```

- Testear a partir de las 100 instancias (sin barajar), uno contra todos para la clase, SVM\_Light.



### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

### Weka (Java)

### Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

### Ejercicios Adicionales

### Referencias

## Ejemplos (Otros algoritmos de minería de datos)

- Barajar datos de entrada, reducción n<sup>o</sup> de atributos a 15, validación cruzada 3 particiones con el clasificador 3NN.

```
Arsenal('classify -t DataExample1.txt -sf 1 -svd 15 -drf DataExample1svd.txt --cross_validate -t 3 --kNN_classify -k 3');
```

- Testear a partir de las 100 instancias (sin barajar), uno contra todos para la clase, SVM\_Light.

```
Arsenal('classify -t DataExample2.txt -sf 0 -- Train_Test_Validate -t 100 -- Train_Test_Multiple_Class -- SVM_Light -Kernel 2 -KernelParam 0.01 -CostFactor 3');
```



### MATLAB Arsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLAB Arsenal

Opciones

Ejemplos

### Weka (Java)

### Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

### Ejercicios Adicionales

### Referencias

## Ejemplos (Otros algoritmos de minería de datos)

- Barajar datos de entrada, reducción n° de atributos a 15, validación cruzada 3 particiones con el clasificador 3NN.

```
Arsenal('classify -t DataExample1.txt -sf 1 -svd 15 -drf DataExample1svd.txt --cross_validate -t 3 --kNN_classify -k 3');
```

- Testear a partir de las 100 instancias (sin barajar), uno contra todos para la clase, SVM\_Light.

```
Arsenal('classify -t DataExample2.txt -sf 0 -- Train_Test_Validate -t 100 -- Train_Test_Multiple_Class -- SVM_Light -Kernel 2 -KernelParam 0.01 -CostFactor 3');
```

- Especificar archivo de test, Boosting con C4.5 de Weka.



### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

### Weka (Java)

### Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

### Ejercicios Adicionales

### Referencias

## Ejemplos (Otros algoritmos de minería de datos)

- Barajar datos de entrada, reducción n° de atributos a 15, validación cruzada 3 particiones con el clasificador 3NN.

```
Arsenal('classify -t DataExample1.txt -sf 1 -svd 15 -drf DataExample1svd.txt --cross_validate -t 3 --kNN_classify -k 3');
```

- Testear a partir de las 100 instancias (sin barajar), uno contra todos para la clase, SVM\_Light.

```
Arsenal('classify -t DataExample2.txt -sf 0 -- Train_Test_Validate -t 100 -- Train_Test_Multiple_Class -- SVM_Light -Kernel 2 -KernelParam 0.01 -CostFactor 3');
```

- Especificar archivo de test, Boosting con C4.5 de Weka.

```
Arsenal('classify -t DataExample2.train.txt -- Test_File_Validate -t DataExample2.test.txt -- MCAdaBoostM1 -- WekaClassify -MultiClassWrapper 0 -- trees.J48');
```



## Ejemplos (Otros algoritmos de minería de datos)

- Barajar datos de entrada, reducción n° de atributos a 15, validación cruzada 3 particiones con el clasificador 3NN.

```
Arsenal('classify -t DataExample1.txt -sf 1 -svd 15 -drf DataExample1svd.txt --cross_validate -t 3 --kNN_classify -k 3');
```

- Testear a partir de las 100 instancias (sin barajar), uno contra todos para la clase, SVM\_Light.

```
Arsenal('classify -t DataExample2.txt -sf 0 -- Train_Test_Validate -t 100 -- Train_Test_Multiple_Class -- SVM_Light -Kernel 2 -KernelParam 0.01 -CostFactor 3');
```

- Especificar archivo de test, Boosting con C4.5 de Weka.

```
Arsenal('classify -t DataExample2.train.txt -- Test_File_Validate -t DataExample2.test.txt -- MCAdaBoostM1 -- WekaClassify -MultiClassWrapper 0 -- trees.J48');
```

- Sobrecribir resultado anterior, 3 clasificadores Naive Bayes con (1.,120), (121.,150) y (154.,225). Seleccionar mayoría por voto.



## Ejemplos (Otros algoritmos de minería de datos)

- Barajar datos de entrada, reducción n° de atributos a 15, validación cruzada 3 particiones con el clasificador 3NN.

```
Arsenal('classify -t DataExample1.txt -sf 1 -svd 15 -drf DataExample1svd.txt --cross_validate -t 3 --kNN_classify -k 3');
```

- Testear a partir de las 100 instancias (sin barajar), uno contra todos para la clase, SVM\_Light.

```
Arsenal('classify -t DataExample2.txt -sf 0 -- Train_Test_Validate -t 100 -- Train_Test_Multiple_Class -- SVM_Light -Kernel 2 -KernelParam 0.01 -CostFactor 3');
```

- Especificar archivo de test, Boosting con C4.5 de Weka.

```
Arsenal('classify -t DataExample2.train.txt -- Test_File_Validate -t DataExample2.test.txt -- MCAdaBoostM1 -- WekaClassify -MultiClassWrapper 0 -- trees.J48');
```

- Sobreescribir resultado anterior, 3 clasificadores Naive Bayes con (1.,120), (121.,150) y (154.,225). Seleccionar mayoría por voto.

```
Arsenal('classify -t DataExample2.txt -oflag w -- Train_Test_Validate -t 100 -- MCWithMultiFSet -Voting -Separator 1,120,121,150,154,225 -- NaiveBayes');
```



## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias



- Algoritmos de aprendizaje automático para tareas de minerías de datos [WF99].
- Desarrollado en la universidad de Waikato, en Nueva Zelanda.
- Herramientas de **preprocesamiento, clasificación, regresión, clustering, reglas de asociación y visualización**.
- Preparado para la inclusión de nuevos algoritmos.
- Código abierto bajo licencia GNU.
- Mucho soporte y documentación:
  - <http://weka.wikispaces.com/>
  - Lista de correo de Weka muy activa.
  - Descripción del paquete de redes Bayesianas: <http://www.metaemotion.com/diego.garcia.morate/>
  - Tutorial Weka en español: <http://www.metaemotion.com/diego.garcia.morate/>
  - Última versión estable: 3.6; versión de desarrollador: 3.7.

Ana M. Martínez



### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

### Weka (Java)

### Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

### Ejercicios Adicionales

### Referencias



## MATLAB Arsenal

- Introducción
- Pasos de Instalación
- Cómo utilizar MATLAB Arsenal
- Opciones
- Ejemplos

## Weka (Java)

## Bayes Net Toolbox

- Introducción
- Ejercicio
- Paquete de aprendizaje estructural de Leray y François
- Ejemplos

## Ejercicios Adicionales

## Referencias

Preprocess   Classify   Cluster   Associate   Select attributes   Visualize

Open fil...   Open U...   Open D...   Genera...   Undo   Edit...   Save...

Filter

Choose   None   Apply

Current relation

Relation: weather   Attributes: 5  
Instances: 14   Sum of weights: 14

Attributes

All   None   Invert   Pattern

| No. | Name        |
|-----|-------------|
| 1   | outlook     |
| 2   | temperature |
| 3   | humidity    |
| 4   | wind-speed  |

Remove

Selected attribute

Name: outlook   Type: Nominal  
Missing: 0 (0%)   Distinct: 3   Unique: 0 (0%)

| No. | Label    | Count | Weight |
|-----|----------|-------|--------|
| 1   | sunny    | 5     | 5,0    |
| 2   | overcast | 4     | 4,0    |
| 3   | rain     | 5     | 5,0    |

Class: play (Nom)   Visualize All

Status

OK   Log   x 0

Ana M. Martínez



## MATLABArsenal

- Introducción
- Pasos de Instalación
- Cómo utilizar MATLABArsenal
- Opciones
- Ejemplos

## Weka (Java)

### Bayes Net Toolbox

- Introducción
- Ejercicio
- Paquete de aprendizaje estructural de Leray y François
- Ejemplos

## Ejercicios Adicionales

## Referencias

Setup Run Analyse

Experiment Configuration Mode:  Simple  Advanced

Open... Save... New

Results Destination

ARFF file  Filename:  Browse...

Experiment Type

Cross-validation

Number of folds: 10

Classification  Regression

Iteration Control

Number of repetitions: 10

Data sets first  Algorithms first

Datasets

Add new... Edit select... Delete sel...

Use relative ...

/home/ana/weka-3-7-0/data/weather.arff

Up Down

Algorithms

Add ne... Edit select... Delete select...

NaiveBayes

Load option... Save option... ... Do...

Notes

Ana M. Martínez



DataSources DataSinks Filters Classifiers Clusters Associations Evaluation Visualization

ADD BayesianLogistic Bayes Complement DMN HNB Naive NaiveBayes NaiveBayes NaiveBayes NaiveBayes NaiveBayes

Esr Regression Net NaiveBayes Btext HNB Bayes Multinomial MultinomialUpdateable Simple Updateable

Knowledge Flow Layout

```

    graph LR
      A[ARFF Loader] -- dataSet --> B[Discretize]
      B -- dataSet --> C[CrossValidation FoldMaker]
      C -- trainingSet --> D[NaiveBayes]
      C -- testSet --> E[NaiveBayes]
      D -- text --> F[TextViewer]
      E -- text --> F
  
```

Status Log

| Component       | Parameters | Time  | Status                             |
|-----------------|------------|-------|------------------------------------|
| [KnowledgeFlow] |            | 0:8:4 | Welcome to the Weka Knowledge Flow |

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje

estructural de Leray y

François

Ejemplos

## Ejercicios Adicionales

## Referencias

# Características principales en el entorno de las redes Bayesianas

- Clasificadores Bayesianos: **Naive Bayes**, **AODE**, **TAN**, otros clasificadores para texto, etc.
- **Aprendizaje estructural** utilizando varios Hill Climbing (K2,B, etc.) y algoritmos de propósito general (simulated annealing, búsqueda tabú).
- **Métricas locales** implementadas: Bayes, BDe, MDL, entropía y AIC.
- **Métricas globales** implementadas: validación cruzada: leave one out y acumulativa también.
- **Independencia condicional** basada en algoritmos de recuperación causal.
- **Estimación de parámetros** utilizando estimaciones directas y promediado de modelos Bayesianos.
- **Visor de redes Bayesianas.**



## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

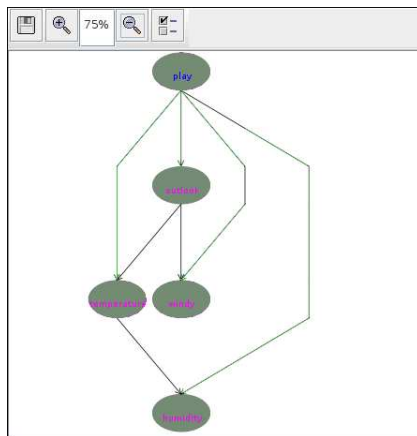
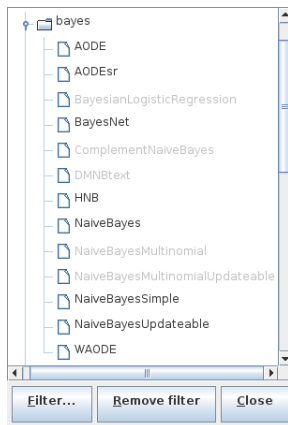
Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## Ejercicios Adicionales

## Referencias



## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje

estructural de Leray y

François

Ejemplos

## Ejercicios Adicionales

## Referencias

## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias



Ana M. Martínez



- BNT de K. Murphy [Mur] no directamente orientado a clasificación.
- Sencillo definir nuevos clasificadores más o menos eficientes.
  - Definir la **variable clase** como **nodo no observado**.
- Poca información, ayuda a través de:  
'help nombre\_función'.
- Falta sistema de evaluación general.

MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje

estructural de Leray y

François

Ejemplos

Ejercicios Adicionales

Referencias

## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

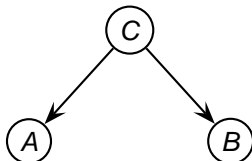
Ejercicios Adicionales

Referencias

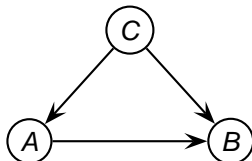
## Ejercicio:

- 1 Crear un clasificador Naive Bayes con 2 atributos discretos ( $A$ ,  $B$ ) y la clase ( $C$ ). CPTs de forma aleatoria ( $C1$ ).
- 2 Crear otro clasificador igual al anterior con un enlace extra entre los dos atributos ( $A \rightarrow B$ ) ( $C2$ ).
- 3 Obtener una base de datos representativa de  $C1$  (muestreo) y utilizarla para obtener los parámetros de  $C1$  y  $C2$ .
  - Pista: `sample_bnet(bnet)`
- 4 Obtener la etiqueta de salida para cada clasificador en todas las combinaciones de atributos.
- 5 Extraer un conjunto de test adicional y testear los dos clasificadores para obtener su tasa de error.
- 6 Repetir los pasos 3 – 5 con  $C2$  como modelo para generar los datos y entrenar los parámetros de  $C1$  y  $C2$ .





(a) Clasificador 1 (NB)



(b) Clasificador 2 (TAN)

**Figura:** Estructuras ejemplo.

| $P(C)$ |     |
|--------|-----|
| $c_1$  | 0,5 |
| $c_2$  | 0,5 |

(a)

| $P(A C)$ | $c_1$ | $c_2$ |
|----------|-------|-------|
| $a_1$    | 0,3   | 0,2   |
| $a_2$    | 0,4   | 0,6   |
| $a_3$    | 0,3   | 0,2   |

(b)

| $P(B C)$ | $c_1$ | $c_2$ |
|----------|-------|-------|
| $b_1$    | 0,4   | 0,5   |
| $b_2$    | 0,6   | 0,5   |

(c)

| $P(B C)$ | $c_1$ |       |       | $c_2$ |       |       |
|----------|-------|-------|-------|-------|-------|-------|
|          | $a_1$ | $a_2$ | $a_3$ | $a_1$ | $a_2$ | $a_3$ |
| $b_1$    | 0,5   | 0,6   | 0,7   | 0,3   | 0,4   | 0,1   |
| $b_2$    | 0,5   | 0,4   | 0,3   | 0,7   | 0,6   | 0,9   |

(d)

**Figura:** Tablas de probabilidad condicionadas.

Ana M. Martínez



## MATLABArsenal

- Introducción
- Pasos de Instalación
- Cómo utilizar MATLABArsenal
- Opciones
- Ejemplos

## Weka (Java)

## Bayes Net Toolbox

- Introducción

### Ejercicio

- Paquete de aprendizaje estructural de Leray y François
- Ejemplos

## Ejercicios Adicionales

## Referencias

## Creación clasificadores

```
N=3; C=1;A=2;B=3;
node_sizes = [2 3 2];
onodes = [2 3];
% CLASIFICADOR 1 (NAIVE BAYES)
dag1 = false(N,N);
dag1(C,[A,B])=true;
bnet1 = mk_bnet(dag1,node_sizes, 'names',{ 'C','A','B'},
                'observed',onodes);
bnet1.CPD{C} = tabular_CPD(bnet1,C,[0.5 0.5]);
bnet1.CPD{A} = tabular_CPD(bnet1,A,[0.3 0.2 0.4 0.6 0.3 0.2]);
bnet1.CPD{B} = tabular_CPD(bnet1,B,[0.4 0.5 0.6 0.5]);
disp('Parámetros NB'); showCPT(bnet1);
% CLASIFICADOR 2
dag2 = false(N,N);
dag2(C,[A,B])=true;dag2(A,B) = true;
bnet2 = mk_bnet(dag2,node_sizes, 'names',{ 'C','A','B'},
                'observed',onodes);
bnet2.CPD{C} = tabular_CPD(bnet2,C,[0.5 0.5]);
bnet2.CPD{A} = tabular_CPD(bnet2,A,[0.3 0.2 0.4 0.6 0.3 0.2]);
bnet2.CPD{B} = tabular_CPD(bnet2,B,[0.5 0.6 0.7 0.3 0.4 0.1
                                     0.5 0.4 0.3 0.7 0.6 0.9]);
disp('Parámetros NB con enlace extra'); showCPT(bnet2);
subplot(1,2,1),draw_graph(dag1);title('Clasificador 1');
subplot(1,2,2),draw_graph(dag2);title('Clasificador 2');
```



### MATLABArsenal

[Introducción](#)[Pasos de Instalación](#)[Cómo utilizar](#)[MATLABArsenal](#)[Opciones](#)[Ejemplos](#)

### Weka (Java)

### Bayes Net Toolbox

[Introducción](#)[Ejercicio](#)[Paquete de aprendizaje  
estructural de Leray y  
François](#)[Ejemplos](#)

### Ejercicios Adicionales

### Referencias

# Muestreo de C1 y recálculo de parámetros

Clasificación con RBs

Ana M. Martínez



**MATLABArsenal**

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

## Muestreo de C1 y recálculo de parámetros

```
for i=1:1000
    s = sample_bnet(bnet1);
    data(1,i)=s{3};
    data(2,i)=s{2};
    data(3,i)=s{1};
end
% Recálculo de parámetros para C1 y C2
bnet1m = learn_params(bnet1, data);
bnet2m = learn_params(bnet2, data);
% p(c=2|a=1,b=3)?
% En el clasificador 1 (NB)
engine = jtree_inf_engine(bnet1);
evidence = cell(1,N);
evidence{A} = 1; evidence{B} = 2;
[engine, loglik] = enter_evidence(engine, evidence);
marg = marginal_nodes(engine, C);
pc1=marg.T(2)
% En el clasificador 2 (NB)
engine = jtree_inf_engine(bnet2);
evidence = cell(1,N);
evidence{A} = 1; evidence{B} = 2;
[engine, loglik] = enter_evidence(engine, evidence);
marg = marginal_nodes(engine, C);
pc2=marg.T(2)
```



### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

### Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias



# Muestreo adicional de C1 para hallar las tasas de acierto

Clasificación con RBs

Ana M. Martínez



**MATLABArsenal**

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

**Bayes Net Toolbox**

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

## Muestreo adicional de C1 para hallar las tasas de acierto

```
cor = 0;
for i=1:100
    s = sample_bnet(bnet1);
    evidence = s'; evidence{C} = [];
    engine = jtree_inf_engine(bnet1);
    [engine, loglik] = enter_evidence(engine, evidence);
    marg = marginal_nodes(engine, C);
    [pr v]= max(marg.T);
    if v==s{1}
        cor=cor+1;
    end
end
accuracy1 = cor/100
cor = 0;
for i=1:100
    s = sample_bnet(bnet1);
    evidence = s'; evidence{C} = [];
    engine = jtree_inf_engine(bnet2);
    [engine, loglik] = enter_evidence(engine, evidence);
    marg = marginal_nodes(engine, C);
    [pr v]= max(marg.T);
    if v==s{1}
        cor=cor+1;
    end
end
accuracy2 = cor/100
```



- 1 ¿Qué ocurre si en lugar de muestrear los datos del clasificador 1 lo hacemos desde el 2 y repetimos los pasos anteriores?
- 2 ¿Cuál es el efecto de la estructura del modelo en la tasa de acierto o error?
- 3 Trata de explicar el resultado experimental como una función de la estructura de la red y los parámetros aprendidos.
- 4 Dados los parámetros para C2, ¿puedes derivar los parámetros equivalentes para C1? ¿Y al revés?
- 5 ¿Cómo podrías maximizar la diferencia de rendimiento entre los dos modelos?
- 6 Compara la tasa de acierto en los experimentos con el error de Bayes (mínimo posible) para las dos estructuras.
  - $Error_{Bayes} = \frac{1}{|sp(t)|} \sum_{x \in sp(t)} (1 - (p(c_1)p(x|c_1) + p(c_2)p(x|c_2)))$



### MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

### Weka (Java)

### Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje

estructural de Leray y

François

Ejemplos

### Ejercicios Adicionales

### Referencias

## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias



- Clasificadores basado en redes Bayesianas incluidos:
  - 1 **NB.**
  - 2 **TAN** y FAN.
  - 3 MWST (Maximum Weight Spanning Tree).
  - 4 PC (búsqueda causal, tests de independencia).
  - 5 K2 (con diferentes inicializaciones).
  - 6 MCMC (Markov Chain Monte Carlo).
  - 7 GS (Greedy Search).
  - 8 GES (Greedy Equivalent Search).
  - 9 SEM (Structural EM).
- Otros clasificadores:
  - 1 kNN.
- Útil si ya se utiliza el paquete para otras tareas de aprendizaje.
- De nuevo, poca documentación [LF].

## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

## Ejercicios Adicionales

## Referencias

# Ejemplo: Aprendizaje “estructural” y paramétrico de NB con la base de datos weather.data

- Pista: `help mk_naive_struct;`



## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

## Ejercicios Adicionales

## Referencias

## Ejemplo: Aprendizaje “estructural” y paramétrico de NB con la base de datos weather.data

- Pista: `help mk_naive_struct;`

```
clear all; close all;
data=load('weather.data');
class=5;
dag = mk_naive_struct(5, class);
% Con BNT:
% dag = false(5);
% dag(5, [1,2,3,4]) = true;
ns= [3 1 1 2 2];
bnet = mk_bnet(dag, ns, 'discrete', [1 4 5]);
bnet.CPD{5}=tabular_CPD(bnet, 5);
bnet.CPD{1}=tabular_CPD(bnet, 1);
bnet.CPD{2}=gaussian_CPD(bnet,2);
bnet.CPD{3}=gaussian_CPD(bnet,3);
bnet.CPD{4}=tabular_CPD(bnet, 4);
bnet = learn_params(bnet,data');
% ¡Ojo! En classification_evaluation (SLP):
% Línea 13, cambiar rationplus por ratioplus
% Intervalo de confianza del 95%:
[ratio, ratiominus, ratioplus] = classification_evaluation
    (bnet, data', class)
```



# Ejemplo: Aprendizaje estructural y paramétrico de TAN (Información mutua condicionada y BIC) y FAN con weather.data

- Pista: `help learn_struct_tan; help learn_struct_fan2`



## MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

## Ejercicios Adicionales

## Referencias



## Ejemplo: Aprendizaje estructural y paramétrico de TAN (Información mutua condicionada y BIC) y FAN con weather.data

- Pista: `help learn_struct_tan`; `help learn_struct_fan2`

```
clear all; close all;
fprintf('\n=== Clasificadores FAN y TAN con SLP \n');
data=load('weather.data');
ns=max(data); class=5;
tmp=cputime;
dag = learn_struct_fan2(data', class, ns);
fprintf('Tiempo construcción FAN: %f\n',cputime-tmp);
tmp=cputime;
dag2=learn_struct_tan(data', class,1, ns,'mutual_info');
fprintf('Tiempo construcción TAN (inf. mutua):
      %f\n',cputime-tmp);
tmp=cputime;
dag3=learn_struct_tan(data', class,1, ns,'bic');
fprintf('Tiempo construcción TAN (bic): %f\n',cputime-tmp);
figure;
subplot(1,3,1),draw_graph(dag);title('FAN: inf.mutua cond. ');
subplot(1,3,2),draw_graph(dag2);title('TAN: inf.mutua cond. ');
subplot(1,3,3),draw_graph(dag3);title('TAN: medida BIC');
```



### MATLABArsenal

Introducción  
Pasos de Instalación  
Cómo utilizar  
MATLABArsenal  
Opciones  
Ejemplos

### Weka (Java)

### Bayes Net Toolbox

Introducción  
Ejercicio  
Paquete de aprendizaje  
estructural de Leray y  
François

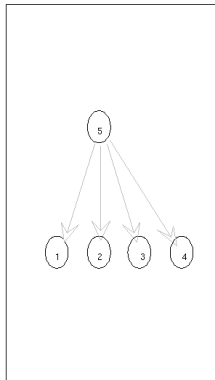
### Ejemplos

### Ejercicios Adicionales

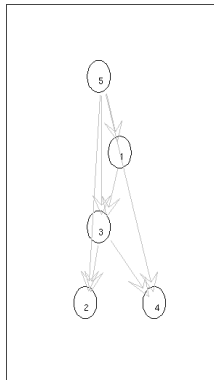
### Referencias



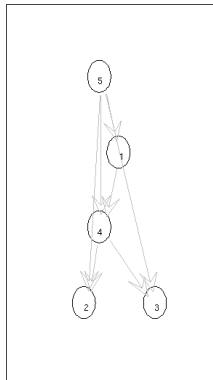
FAN: inf.mutua cond.



TAN: inf.mutua cond.



TAN: medida BIC



## MATLABarsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABarsenal

Opciones

Ejemplos

## Weka (Java)

## Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

## Ejercicios Adicionales

## Referencias

## 1 MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar MATLABArsenal

Opciones

Ejemplos

## 2 Weka (Java)

## 3 Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje estructural de Leray y François

Ejemplos

## 4 Ejercicios Adicionales

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

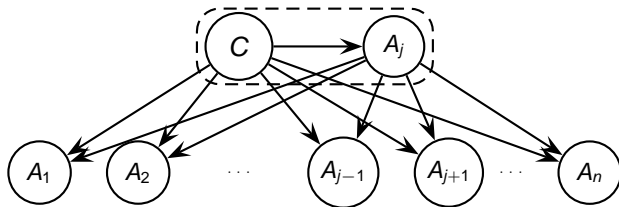
Paquete de aprendizaje  
estructural de Leray y  
François

Ejemplos

Ejercicios Adicionales

Referencias

- 1 Repetir el ejercicio de la transparencia 39 para datos numéricos (gaussianas).
  - Pista: gaussian\_CPD
- 2 Implementar NB con/sin utilizar BNT (discretas y Gaussianas).
- 3 Implementar un clasificador AODE [WBW05].



**Figura:** Estructura clasificador AODE.

**Dudas:** [anamaria.martinez@uclm.es](mailto:anamaria.martinez@uclm.es)



- [ars] **Matlabarsenal.**  
<http://www.informedia.cs.cmu.edu/yanrong/MATLABArsenal/MATLABArsenal.htm>.
- [LF] **Philippe Leray and Olivier François.**  
**Paquete de aprendizaje estructural con bnt.**  
<http://bnt.insa-rouen.fr/ajouts.html>.
- [Mur] **Kevin Murphy.**  
**Bayes net toolbox for matlab.**  
<http://code.google.com/p/bnt/>.
- [WBW05] **G. I. Webb, J. R. Boughton, and Z. Wang.**  
**Not So Naive Bayes: Aggregating One-Dependence Estimators.**  
*Machine Learning*, 58(1):5–24, 2005.



**MATLABArsenal**

- Introducción
- Pasos de Instalación
- Cómo utilizar MATLABArsenal
- Opciones
- Ejemplos

**Weka (Java)**

**Bayes Net Toolbox**

- Introducción
- Ejercicio
- Paquete de aprendizaje estructural de Leray y François
- Ejemplos

**Ejercicios Adicionales**

**Referencias**

- [WF99] I. H. Witten and E. Frank.  
*Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations.*  
Morgan Kaufmann, October 1999.

Ana M. Martínez



MATLABArsenal

Introducción

Pasos de Instalación

Cómo utilizar

MATLABArsenal

Opciones

Ejemplos

Weka (Java)

Bayes Net Toolbox

Introducción

Ejercicio

Paquete de aprendizaje

estructural de Leray y

François

Ejemplos

Ejercicios Adicionales

Referencias