

## PROGRAMAS

AMBIENTAL ☐ BIOINGENIERÍA ☐ ELECTRONICA ☐ INDUSTRIAL ☐ SISTEMAS ☐ OTROS ☐NOMBRE DEL DOCENTE: Mónica Rojas Martínez

IDENTIFICACION:

SEMESTRE: VI

SALON/AULA: E202PERIODO: 2018-1**PRESENTACION/INTRODUCCION**

*El centro de desarrollo tecnológico es una unidad encargada de la prestación de servicios y le recomendamos acatar las normas y ordenamiento para su adecuado uso, y por lo tanto le sugerimos que todos los contenidos y normas las utilice para su adecuado uso.*

En este laboratorio se implementará un modelo de clasificación múltiple del tipo uno contra todos basado en regresión logística. Este tipo de modelo es capaz de definir límites de decisión lineales pero se pueden transformar las entradas para lograr límites más complejos.

Adicionalmente se implementará el factor de Regularización y se analizará su efecto sobre la clasificación.

**CONTENIDO**NOMBRE DE LA PRÁCTICA: Clasificación uno contra todos

PRACTICA NUMERO: VI

**OBJETIVO GENERAL**

Familiarizarse con la clasificación multiclase en el modelo uno contra todos

**OBJETIVOS ESPECIFICOS**

1. Implementar una serie de algoritmos para obtener una clasificación multiclase basada en regresión logística
2. Entrenar un modelo de clasificación uno contra todos
3. Observar el efecto de la variación del factor de regularización

## RESULTADOS ESPERADOS

A partir de la base de datos Iris obtener la clasificación en las diferentes familias

## CONTENIDO/METODOLOGIA

1. Cargar los datos del laboratorio en el archivo iris.mat. Los datos están divididos en un conjunto de entrenamiento ([X, y]) y uno de validación ([Xv, yv])
2. Implementar la función sigmoide correspondiente a la hipótesis de la regresión logística como:

```
function h = hipoLog (z)
```

donde:

$$h(\theta) = \frac{1}{1 + e^{-z}}$$

3. Implementar la función de coste de la regresión logística. Esta función debe devolver tanto el coste como el valor del gradiente de la función de coste. El gradiente servirá para actualizar los parámetros theta. La función de coste debe tener en cuenta un término de regularización para evitar el sobreajuste del algoritmo. La definición de la función en Matlab se debe implementar de la siguiente manera:

```
function [J, grad] = CosteRL(theta, X, y, lambda)
```

4. Implementar la función

```
function [THETA] = UnovsTodosOptim(X, y, num_clases, lambda).
```

Esta función permite entrenar los parámetros theta para asignar la pertenencia a cada una de las clases. Recuerde que el algoritmo Uno contra Todos hace un clasificador para cada una de las clases (ver diapositivas) por lo tanto hay que entrenar un conjunto de parámetros theta para cada clase. En la función, X es la matriz de entrada y y es el vector de clases de pertenencia de la matriz X que contiene tanto la clase positiva como la negativa (recordar que es una clasificación binaria). lambda es el factor de regularización para evitar el sobreajuste del modelo. Recuerde que los parámetros se entrenan ÚNICAMENTE con el conjunto de entrenamiento.

La salida THETA , es una matriz de dimensión número de entradas +1 × número de clases

La actualización de los parámetros se hará con la función fmincg, una función de optimización similar al algoritmo del gradiente descendente pero mucho más eficiente. A continuación se describe como se hace el llamado a esta función:

```
theta= ...  
    fmincg(@(t)(CosteRL(t, X, y, lambda)), theta0, op);
```

La línea “@(t)(CosteRL(t, X, y, lambda))” es una forma de hacer un llamado indirecto a la función. En este caso, t es la variable a ser actualizada (es decir, los parámetros), los demás parámetros se describen a continuación: X es la matriz de entrada; y es el vector de clases de

pertenencia de la matriz X que contiene tanto la clase positiva como la negativa; lambda es el factor de regularización para evitar el sobreajuste del modelo.

Por último theta0 es el vector inicial de theta, y op es un conjunto de opciones que debe ser cargado desde la variable opciones\_opt proveída con los archivos para la realización de este laboratorio.

Se debe cargar de la siguiente manera:

```
load 'opciones_opt.mat';
```

5. Implementar la función

`function` clase = AsignarClase (THETA, X). Esta función asigna la pertenencia de cada una de las observaciones de la matriz X a una de las clases, utilizando los parámetros óptimos THETA. Recuerde que la observación  $x^{(j)}$  pertenece a la clase con mayor probabilidad de salida. Para esta función puede utilizar la función 'max' de Matlab. Para ver su funcionamiento, puede digitar desde la línea de comandos la instrucción `help max`

6. Varíe el parámetro de regularización lambda según los siguientes valores  $\lambda = \{0, 0.5, 1\}$  y calcule el rendimiento del modelo en términos del porcentaje de aciertos (PA) en cada caso.

$$PA = \frac{\# \text{ total observaciones correctamente clasificadas}}{\# \text{ total observaciones}}$$

Que puede concluir acerca del PA para cada uno de los valores lambda?

7. Aunque el porcentaje de aciertos permite obtener una idea global del rendimiento de la clasificación, indique según su criterio que índice(s) es(son) mas apropiado(s) para medir el rendimiento de cada una de las clases? De acuerdo con esto, que índices permitirán tener una mejor idea del rendimiento general de la clasificación? Justifique su respuesta

## MATERIALES, EQUIPOS O RECURSOS

| EQUIPO/MATERIAL       | CANTIDAD | MEDIDA | PLACA/CODIGO |
|-----------------------|----------|--------|--------------|
| Computador con Matlab | 1        |        |              |
|                       |          |        |              |
|                       |          |        |              |
|                       |          |        |              |
|                       |          |        |              |
|                       |          |        |              |
|                       |          |        |              |

## CONDICIONES

- No olvide diligenciar el formato "SOLICITUD DE PRACTICAS DE LABORATORIO", con ocho (8) días de anticipación a la realización de cualquier práctica, para con esto evitar cualquier inconveniente.
- Para solicitar los equipos y materiales anteriormente mencionados, favor diligenciar el formato "SOLICITUD DE MATERIAL DE PRACTICAS DE LABORATORIO".
- En caso de presentarse algún inconveniente o novedad dentro de la realización de las diferentes prácticas, ya sea por parte del profesor o estudiante, diligenciar el formato "REPORTE DE NOVEDADES".