

**INTERNET APPLICATIONS
PROBLEMS-FEBRUARY 2019**

1. We want to extend the video application that has been used for your final project. As you know, the application uses a MySQL server with a DB called *ai*, which stores the table *movies* whose fields are *id*, *title*, *desc*, *url_pic* and *date* with a unique identifier, title, description, URL of the photo and release date. We want to add information and functionality related to the actors in the movies. For this, a table *actors* will be added with fields *id*, *name*, *surname* and *url_foto*, describing an actor. The actors starring in a movie will be stored in a new *moviesactors* table that relates a movie with actors.

Do not just write code, briefly describe your solution.

- a) Provide an HTML form where a user can create a new actor in the system. **(0.5 points).**
 - b) Create a PHP script that receives an actor *id*, and generates an HTML form to associate the actor with a film. To do this, in the form the name and surname of the actor will be displayed and a selection list will be provided with the name of all films in the catalog **(0.5 points).**
 - c) Provide a PHP script to process the form of the previous section, ie associate the actor to a movie in the DB. **(0.5 points).**
 - d) Develop a script in PHP that, given the *id* of two actors using the GET method generates a HTML table with the names of all the movies in which they acted together. **(1 point).**
2. To improve the way in which films are associated with an actor (paragraph b) above) so that a user can select several movies for the same actor, we will use Javascript. For this purpose, an HTML table is displayed with all movies, as in the following HTML code fragment (numbering is included to facilitate the answer if necessary):
1. <P> Show selected films </ p>
 2. <Table>
 3. <Tr> <td> 1 </ td> <td> Toy Story </ td> </ tr>
 4. <Tr> <td> 2 </ td> <td> GoldenEye </ td> </ tr>
 5. ... OTHER ENTRIES
 6. </ Table>

To select the movies of the actor, the user will click in the table rows the movies that wants to associate with the actor. That is, whenever the user clicks in a row the name of the movie will be added to the list. To do this:

- a) **Using only Javascript**, each time a user clicks on any of the rows, the name of the movie will be added to the list of selected movies. The list will be displayed by inserting the name of the movie in a new P element under the one which already exists. Associate properly the Javascript code to the HTML fragment **(1.75 points).**
- b) Provide a stylesheet or Javascript, along with the necessary modifications / additions to the fragment, to make the row selected be displayed in green **(0.75 points).**

3. Consider a movie download application that incorporates a collaborative filtering recommendation algorithm for which $n = 3$ characteristics are defined. We have the following database of scores (defined between 0 and 5):

	Andrés	María	Carlos	Eusebio	Laura
Película 1	4	1	0	0	?
Película 2	3	?	2	?	4
Película 3	?	5	?	4	?
Película 4	2	5	4	?	1

To optimize the cost function, we need the following MATLAB function.

Function `[J, X_grad, Theta_grad] = calculateGradient (X, Theta, Y)`

Where Y is the rating matrix, in which each undefined input contains "-1", the matrix X contains the characteristics of the films, and Θ contains the parameters of the users:

$$X = \begin{bmatrix} \text{---} (x^{(1)})^T \text{---} \\ \text{---} (x^{(2)})^T \text{---} \\ \vdots \\ \text{---} (x^{(n_m)})^T \text{---} \end{bmatrix}, \quad \Theta = \begin{bmatrix} \text{---} (\theta^{(1)})^T \text{---} \\ \text{---} (\theta^{(2)})^T \text{---} \\ \vdots \\ \text{---} (\theta^{(n_u)})^T \text{---} \end{bmatrix}$$

Where n_m is the number of movies and n_u is the number of users. The function returns:

J : a scalar with the value of the cost function for the X Θ matrices provided.

X_{grad} : an $n_m \times n$ matrix with the partial derivatives of the cost function J with respect to each element of X

Θ_{grad} : an $n_u \times n$ matrix with the partial derivatives of the cost function J for each element of Θ

a) Write the necessary MATLAB code to obtain: n_m (num_peliculas), n_u (num_usuarios), n (num_caracteristicas) and the matrix R such that $R(i, j) = 1$ if movie i has been rated by user j and $R(i, j) = 0$ otherwise.

b) Write the required MATLAB code to calculate the cost function J without regularization.

c) Consider that x and Θ are initialized to the following values:

$X = [0.2 \ 0.5 \ 1; \ 0 \ 1 \ 1; \ 1 \ 0 \ 0.5; \ 0 \ 0.5 \ 1];$

$\Theta = [1 \ 0 \ 3; \ 2 \ 1 \ 0; \ 3 \ 2 \ 0; \ 4 \ 1 \ 1; \ 0 \ 2 \ 1];$

Calculate the quadratic error obtained when predicting Andrés' score for movie 4 and predicting Carlos' score for movie 2.

d) Write a code in MATLAB that, given a feature vector x of a movie, allows you to find the most similar movie of all the movies in the matrix X .

e) Suppose we have a set of training data with more movies and more users, and we have about 10000 ratings. We decided to introduce a regularization factor λ into the algorithm. Briefly explain how we could adjust the value of λ to avoid overfitting.

4. Consider the following training data X , y :

$$X = \begin{bmatrix} \text{---} (x^{(1)})^T \text{---} \\ \text{---} (x^{(2)})^T \text{---} \\ \vdots \\ \text{---} (x^{(m)})^T \text{---} \end{bmatrix} \quad \vec{y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{bmatrix}$$

Such that $y^{(i)}$ only takes two possible values $\{0,1\}$. These data will be used to train a logistic regression classifier. The vector of parameters θ of the classifier is given by theta (vector column in Matlab).

(a) Write the code of a function in Matlab, $g = \text{sigmoid}(z)$, to obtain the logistic function value $g(z) = (1 + e^{-z})^{-1}$. In the Matlab function, the argument z can be a scalar, or a vector **(0.5 points)**.

(b) Write a Matlab code to obtain the cost function $J(\theta)$ **(0.5 points)**.

(c) Now suppose data are labeled with 5 possible values $y^{(i)} \in \{1,2,3,4,5\}$. We have 1500 training data, $x^{(i)} \in \mathbb{R}^{15}$. Explain how to perform the multi-class classification, how many parameter vectors ($\theta^{(j)}$) will we obtain and what will the dimension of each vector ($\theta^{(j)}$) be. **(0.5 points)**.

(d) If we want to add the quadratic factors $x_1^2, x_2^2, \dots, x_{15}^2$ and the cubic factors $x_1^3, x_2^3, \dots, x_{15}^3$ to perform the classification, indicate the size of the parameter vectors ($\theta^{(j)}$) obtained **(0.5 points)**.

(e) Suppose you have already trained the multiclass classifier. Therefore, it has N vectors $\theta^{(j)}$ distributed in the rows of the matrix Theta (so that $\theta^{(1)}$ is given by `Theta(1, :)`, $\theta^{(2)}$ is given by `Theta(2, :)`, etc.). Write the code needed to determine which class to assign a new input data x . **(0.5 points)**.