**Nombre:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Fecha: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Instrucciones.** *Lee cuidadosamente cada pregunta y contesta lo que se te pide. No olvides escribir subir todos los archivos .m elaborados para cada inciso que requiera de uso de software y pegar el código de programa identado y comentado. De acuerdo al número de lista en el archivo de calificaciones contestar el problema correspondiente.*

1. (**10 puntos**) Realiza un cuadro comparativo entre el perceptrón y el perceptrón multicapa en donde se indique que utilidad tiene cada uno y en qué se diferencian.
2. (**10 puntos**) De tu expediente toma los últimos 3 dígitos que sean diferentes de cero y de uno. Por ejemplo, del expediente 699128 se tomarían los dígitos 928. De los tres dígitos seleccionados, considere al primero como el número entradas a la red, al intermedio como el número de neuronas en la capa oculta y al tercero como el número de neuronas en la capa de salida. Determine cuántos pesos tiene la red neuronal.

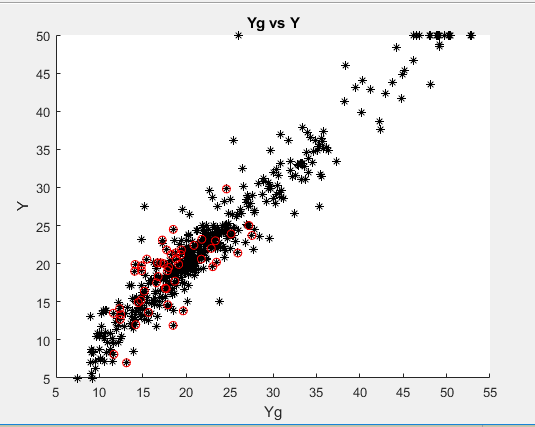
5 9 7

(5+1)\*9 + (9+1)\*7

1. **(20 puntos)** Considere la base de datos contenida en el archivo **house.mat**. Los renglones del 1 al 13 contienen información estadística de 506 vecindarios que será usada para estimar el precio medio de las casas en miles de dólares (Renglón 14).

Las 13 variables (respetando el idioma original son) son:

1. Per capita crime rate per town
2. Proportion of residential land zoned for lots over 25,000 sq. ft.
3. Proportion of non-retail business acres per town
4. 1 if tract bounds Charles river, 0 otherwise.
5. Nitric oxides concentration (parts per 10 million)
6. Average number of rooms per dwelling
7. Proportion of owner-occupied units built prior to 1940
8. Weighted distances to five Boston employment centers
9. Index of accessibility to radial highways
10. Full-value property-tax rate per $10,000
11. Pupil-teacher ratio by town
12. 1000(Bk - 0.63)^2, where Bk is the proportion of blacks by town
13. Percent lower status of the population
14. Utilice el 90% de los datos para entrenar el modelo neuronal y estime el 10 % restante



Los puntos negros son los precios estimados vs los reales en toda la base de datos. Los que están circulados en rojo son los de prueba. Se obtuvo un performance de 7.05 en el entrenamiento y de 10.8 en las pruebas.

1. Estime el precio medio en miles de dólares de una casa en la cual se tienen los siguientes parámetros:

X1= 12.54, X2=45, X3=15.37, X4=1, X5=0.5150, X6=6.1621, X7=45.800, X8=3.3751, X9=7, X10=193, X11=15.200, X12=347.88, X13=2.96

El precio es de aproximadamente 9.22 mil dolares

1. **(40 puntos)** Cargue la base de datos con nombre **ex\_mia4\_data1**. En la base encontrará una matriz de nombre train, con ejemplos de datos que ya fueron clasificados, con dichos datos deberá entrenar una red neuronal para llevar a cabo una clasificación supervisada para categorizar los datos contenidos en la matriz **test\_unknown**.
2. **(5 puntos)** Es altamente probable que su primer intento no funcione. ¿Cuál es el error que impide clasificar los datos y cómo se corrige?

tenemos una variable ‘train’, cuando se manda a llamar la función train el sistema confunde la función con los datos y genera un error. Para evitar este error se debe cambiar el nombre de la variable ‘train’

1. **(15 puntos)** Usando únicamente una neurona en la capa de salida proporcione la clasificación de los datos de la matriz **test\_unknown**.
2. **(15 puntos)** Usando tres neuronas en la capa de salida proporciones la clasificación de los datos de la matriz **test\_unknown**.
3. **(5 puntos)** Compare las clasificaciones obtenidas en el inciso b) y c). ¿Se le asigna la misma clasificación por ambos métodos? Justifique su respuesta.
4. **(20 puntos)** Considere la base de datos proporcionada en el archivo **cancer\_de\_Mama.mat**.

Dicha base contiene la información de una muestra de 699 biopsias. Cada columna representa una variable, las cuales significan lo siguiente:

1. Espesor del grupo.
2. Uniformidad del tamaño de la célula.
3. Uniformidad de la forma de la célula.
4. Adhesión marginal.
5. Tamaño de célula epitelial única.
6. Núcleos desnudos
7. Cromatina suave
8. Núcleo normal
9. Mitosis
10. **(5 puntos)** ¿Los datos en dichas variables requieren de limpieza o tratamiento? Justifique.
11. **(5 puntos)** Determinar La cantidad de grupos óptima en los que conviene agrupar dichos datos. Indicar qué pruebas se determinaron para ellos.
12. **(5 puntos)** Cuántos datos hay en cada grupo.
13. **(5 puntos)** ¿Cuál podría ser la interpretación de cada grupo?