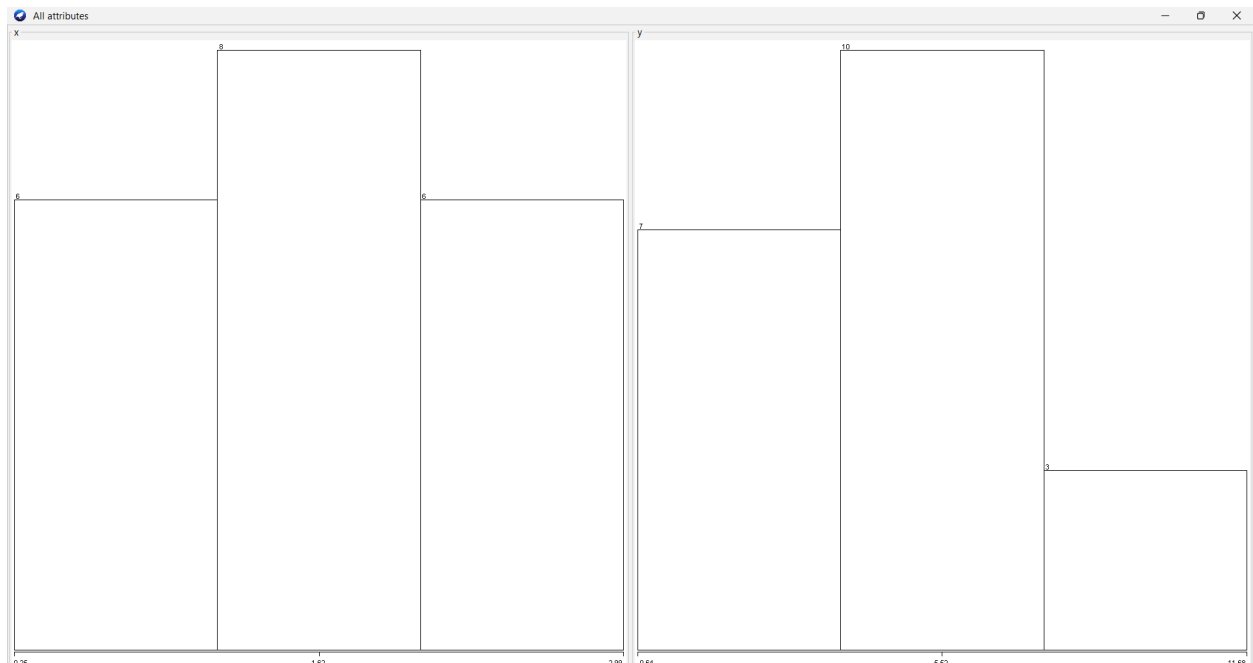
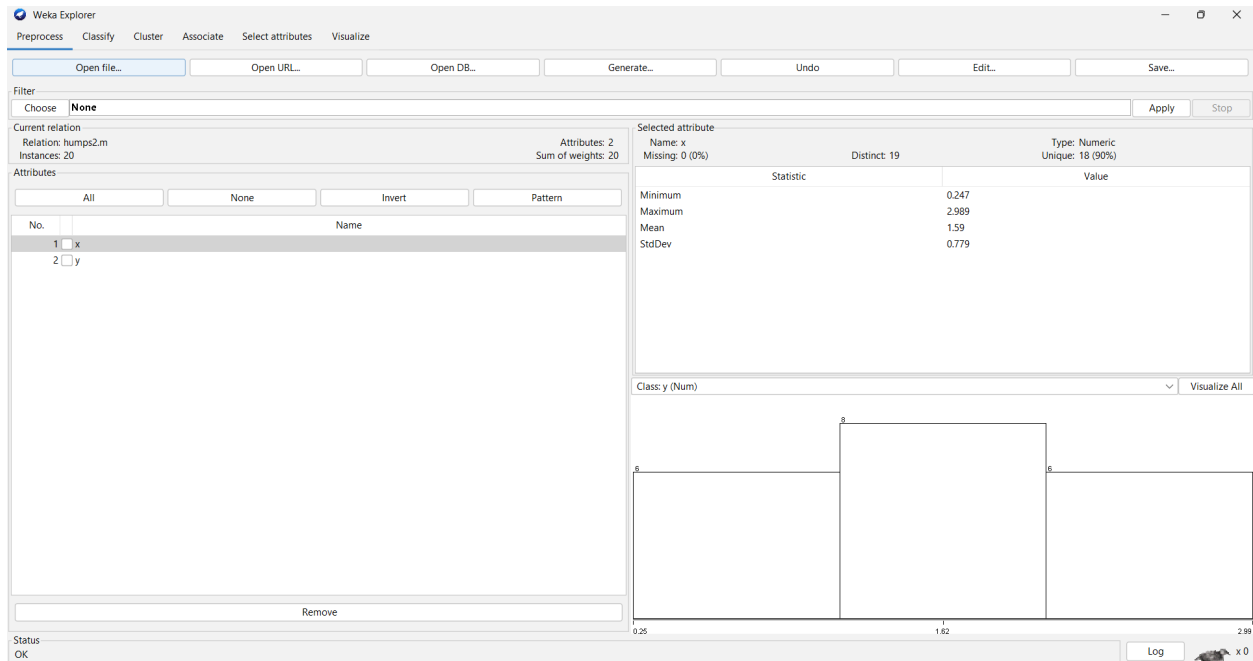


# Practicas 1,2 y 3

## 1. Analisis y descripcion del dataset

Para llevar a cabo este informe usaremos los datasets encontrados en el campus, los cuales se componen de un dataset principal, un entrenamiento y un test. El dataset de entrenamiento se compone de 20 instancias con 2 atributos.



## 2.Entrenamiento del dataset

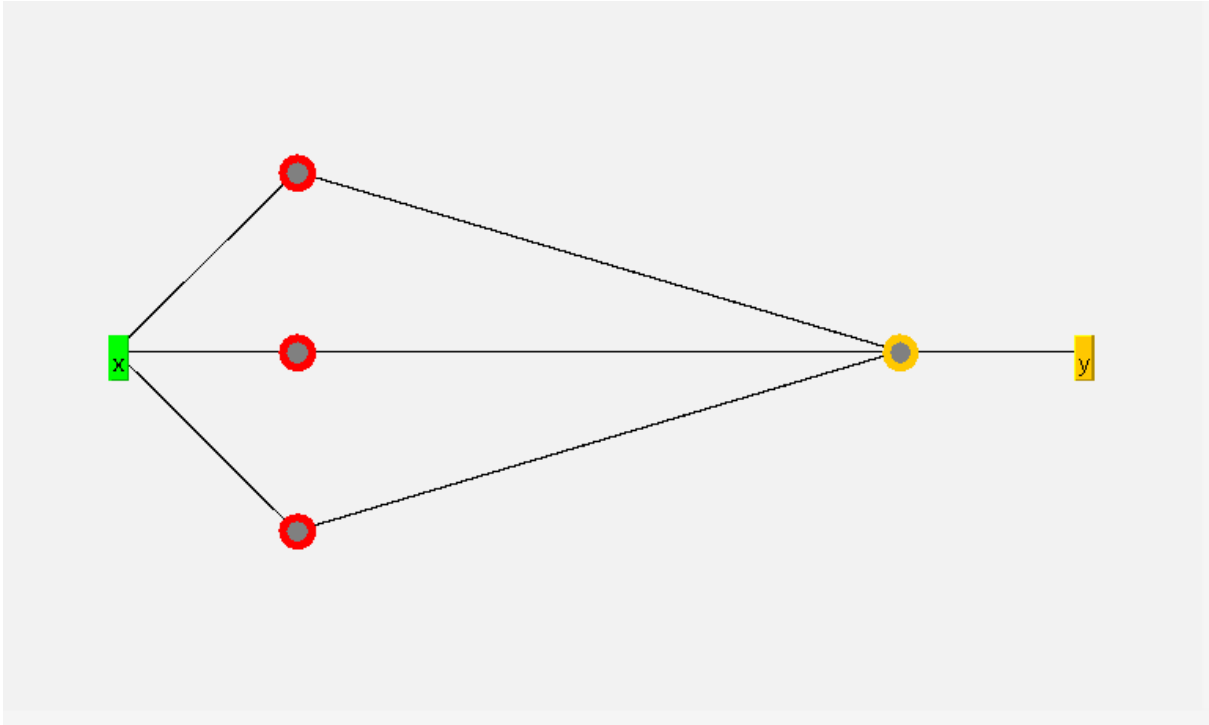
En cada celda el primer valor corresponde al Root Mean Square Error, y el segundo al Relative Square Error

NºCapas Ocultas	Training Set-2000epoch	Test Set	Test Set Config	Cross Validation	Cross Valifation Config
0	2.9946 101.4034 %	1.2595 86.56 %	1.2562 86.3318 %	3.7417 118.7557 %	3.7292 118.3581 %
1	2.9919 101.3116 %	0.7402 50.8709 %	1.5819 108.7174 %	3.4969 110.9865 %	3.0362 96.3652 %
2	2.9672 100.4764 %	0.351 24.1217 %	0.4896 33.6474 %	2.4988 79.3082 %	2.3746 75.3654 %
3	2.9555 100.0796 %	0.3523 24.215 %	0.4845 33.2967 %	1.7131 54.372 %	2.2375 71.0151 %
4	3.0101 101.9282 %	0.4163 28.6116 %	0.5081 34.9179 %	1.7527 55.6275 %	2.2866 72.5748 %
5	3.0164 102.143 %	0.4993 34.3174 %	0.5129 35.2464 %	2.0383 64.6926 %	2.601 82.5528 %
6	3.0181 102.199 %	0.4274 29.3751 %	0.5062 34.7898 %	1.8865 59.8761 %	2.5303 80.3093 %

2 a y b

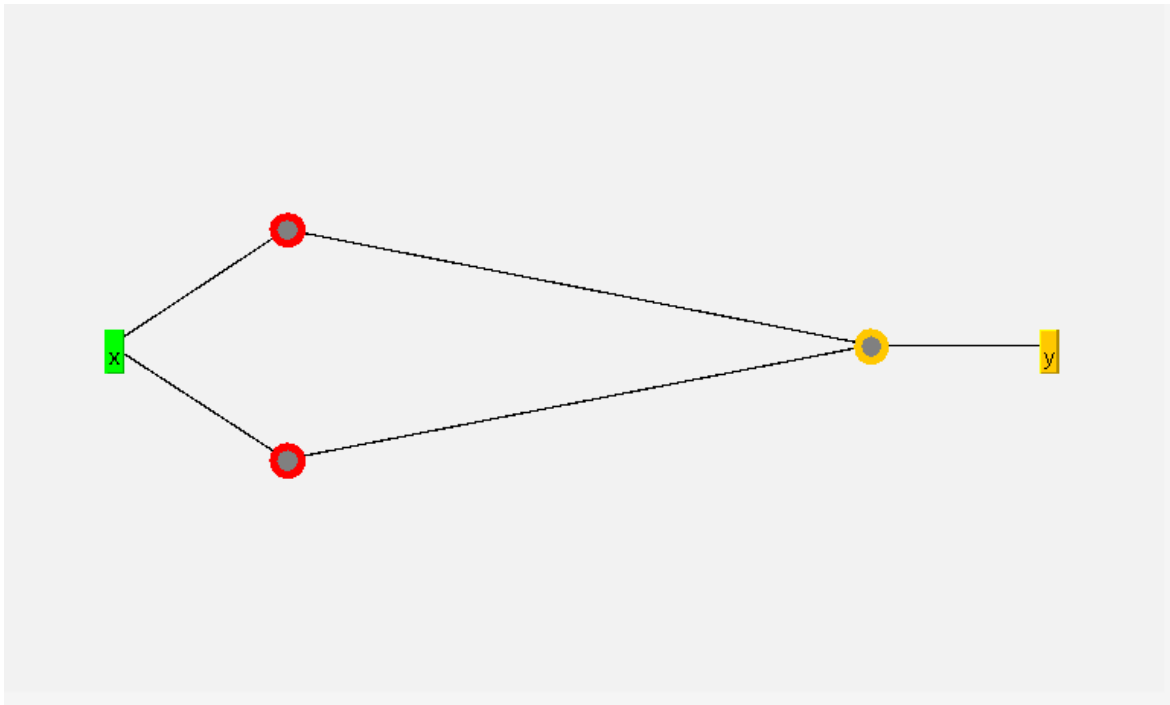
Training Set:

La mejor configuracion con el training set es 3 neuronas en la capa ocultas debido a que el indice de error es minimo, con 2 seria un subajuste y con 4 un sobreajuste.



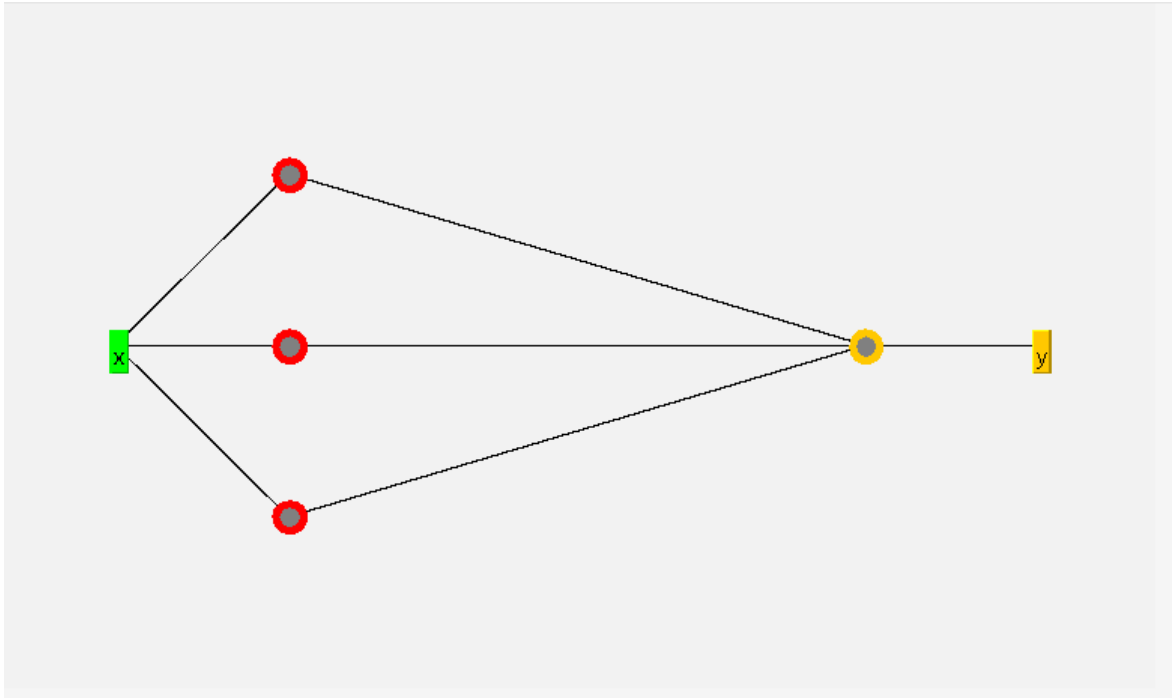
Test Set:

Para el test set los errores minimos son con dos neuronas en la capa oculta siendo esta la mejor opcion



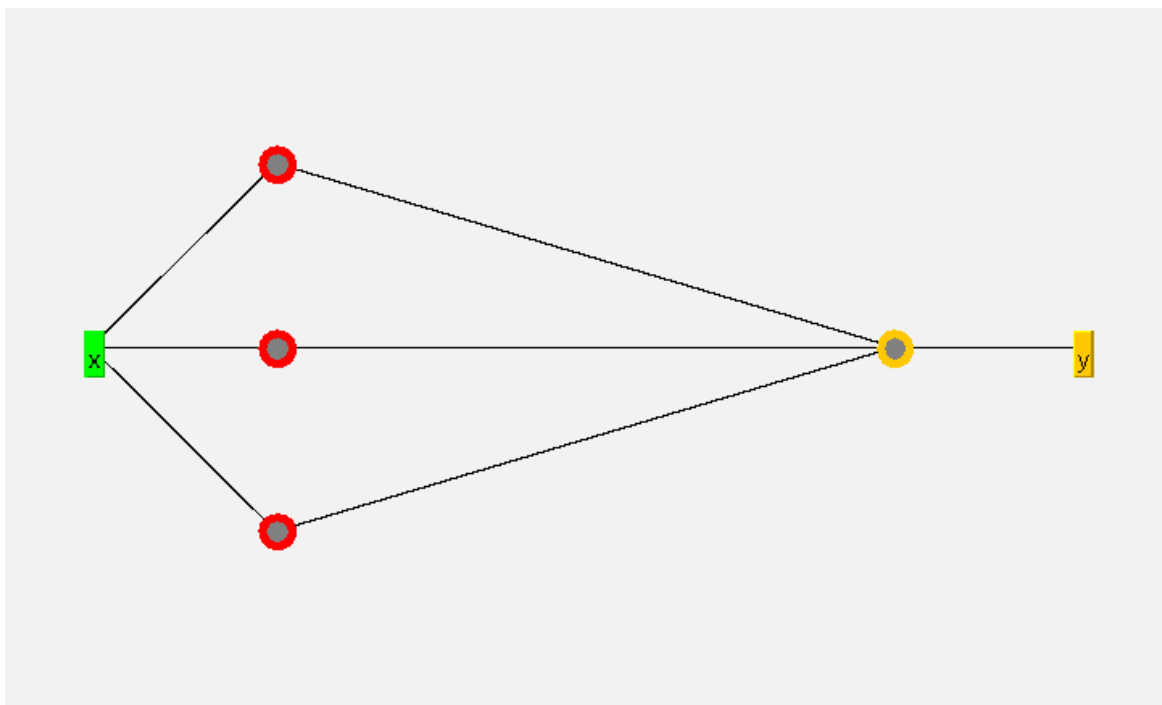
#### Test Set Config:

En el test set con la configuración dada los errores se reducen lo máximo posible con 3 neuronas en la capa oculta



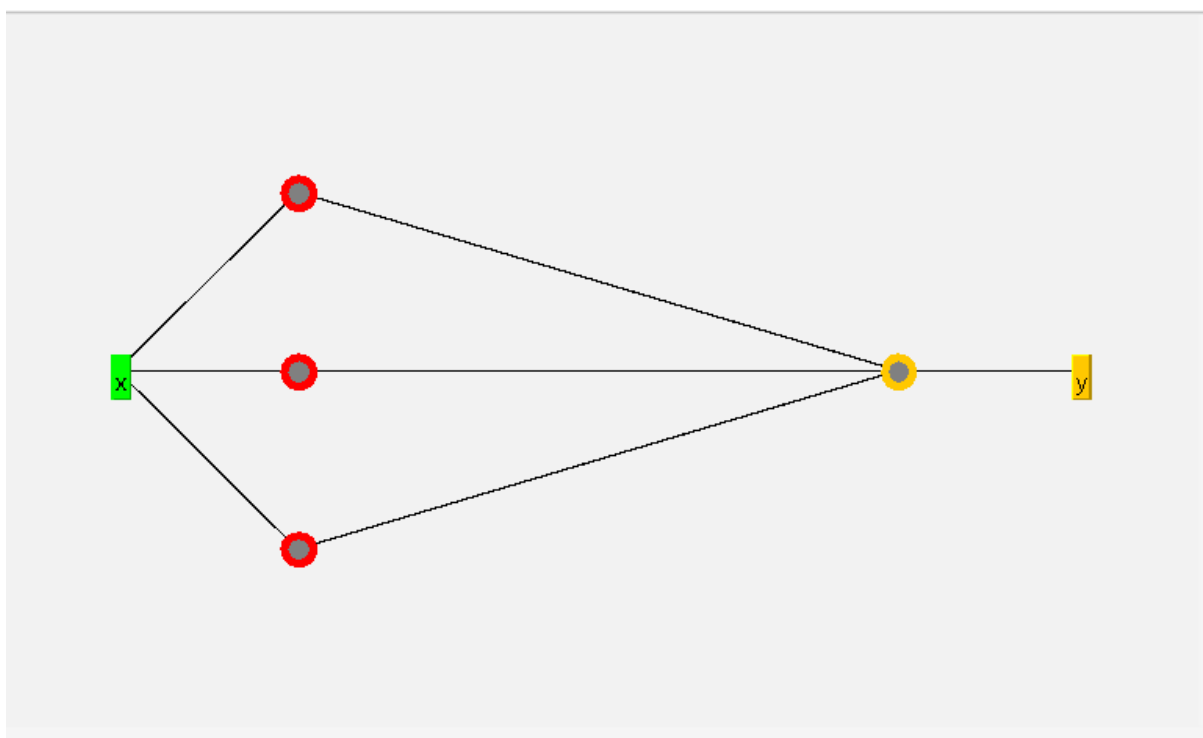
#### Cross Validation:

La mejor configuración para la cross validation sin configurar es con 3 neuronas en la capa oculta dando el error mas pequeño que el resto de configuraciones.



Cross Validation Config:

La mejor configuracion con estas características es 3 neuronas en la capa ocultas debido a que el índice de error es mínimo, con 2 sería un subajuste y con 4 un sobreajuste.



3.a Sí, en cada columna los casos por encima del marcado como amarillo son casos de subajuste ya que no hay la suficiente cantidad de neuronas para que la red este en su funcionamiento mas óptimo

3.b Por el contrario los casos por debajo del marcado en la columna son casos de sobreajuste donde la red se vuelve demasiado concreta por usar mas neuronas de las realmente necesarias

3.c La mejor red neuronal de todas es la del test set ya que da los errores mas bajos aunque tambien es cierto que es la que menos ejemplos de prueba tiene por lo que puede ser algo impreciso

# Memoria Prácticas 3 y 4

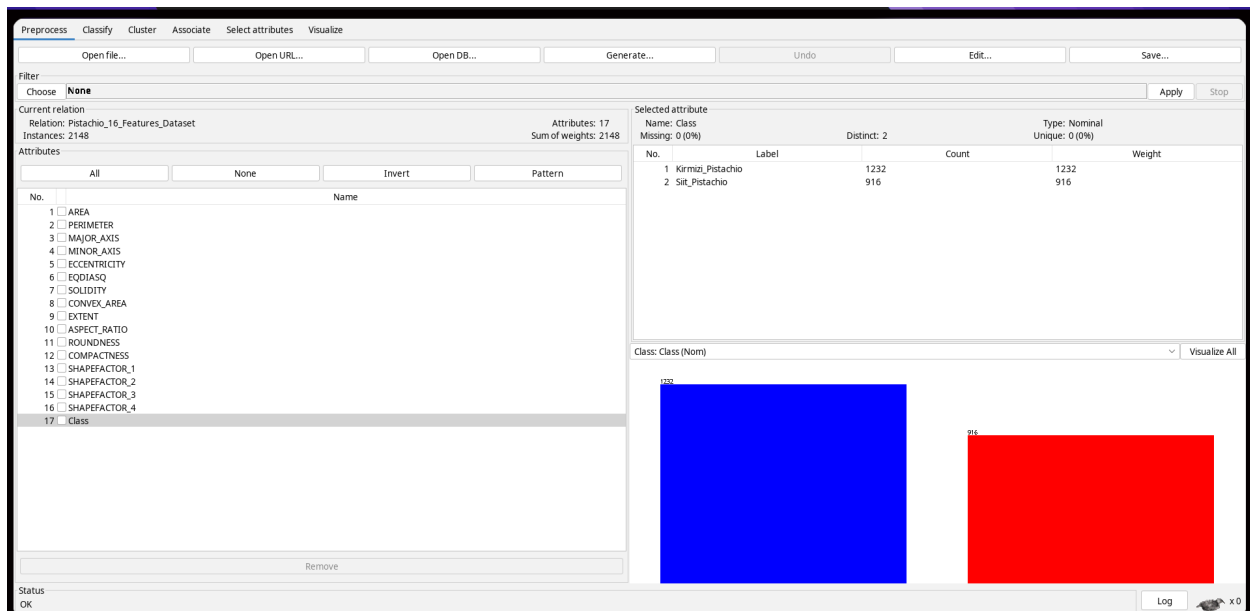
## 1. Elección y descripción del dataset.

Para llevar a cabo esta práctica, hemos escogido un conjunto de datos de la página web <https://www.kaggle.com/> y, concretamente el siguiente:

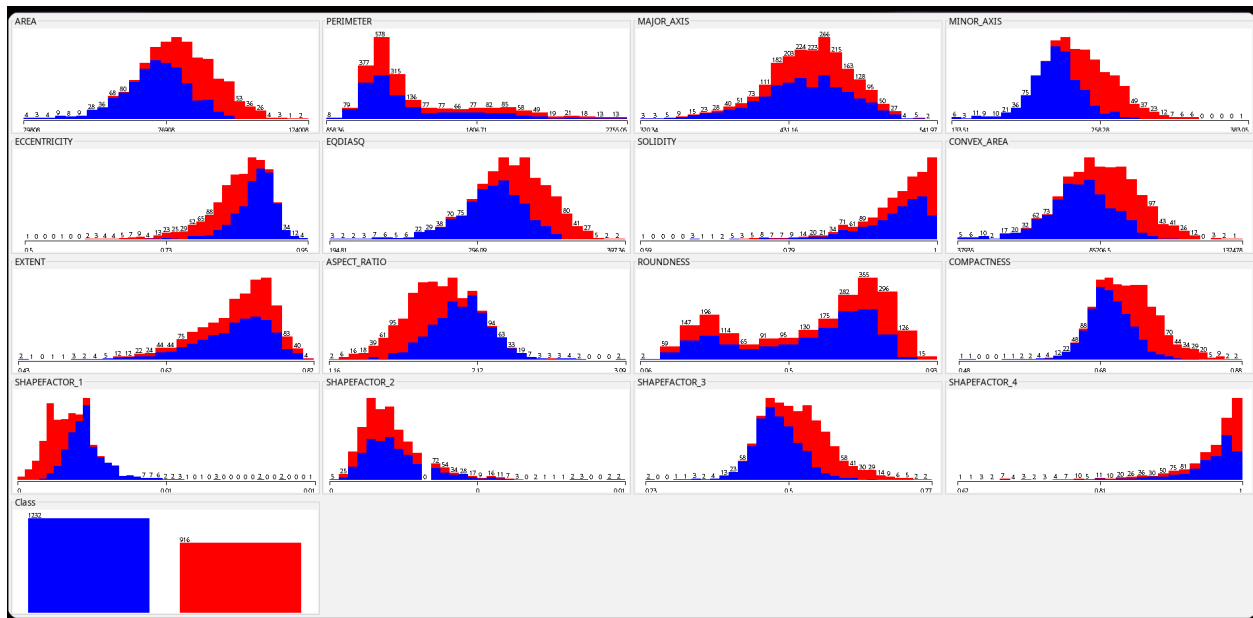
<https://www.kaggle.com/datasets/muratkokludataset/pistachio-dataset>

El dataset está en formato arff y, por lo tanto, ya está preparado para analizarse con Weka. De dicho dataset hemos hecho una partición, donde el 20% de los datos se destina al conjunto de prueba y el 80% restante se destina al conjunto de entrenamiento.

El dataset elegido tiene 2148 instancias con 17 atributos y donde el objetivo es clasificar en dos tipos de pistachos: Kirmizi y Siit, en este caso, hay 1232 de Kirmizi y 916 de Siit.



Información principal del dataset y visualización de los dos tipos de clase.



Visualización de las gráficas para cada atributo.

## 2. Entrenamiento del dataset.

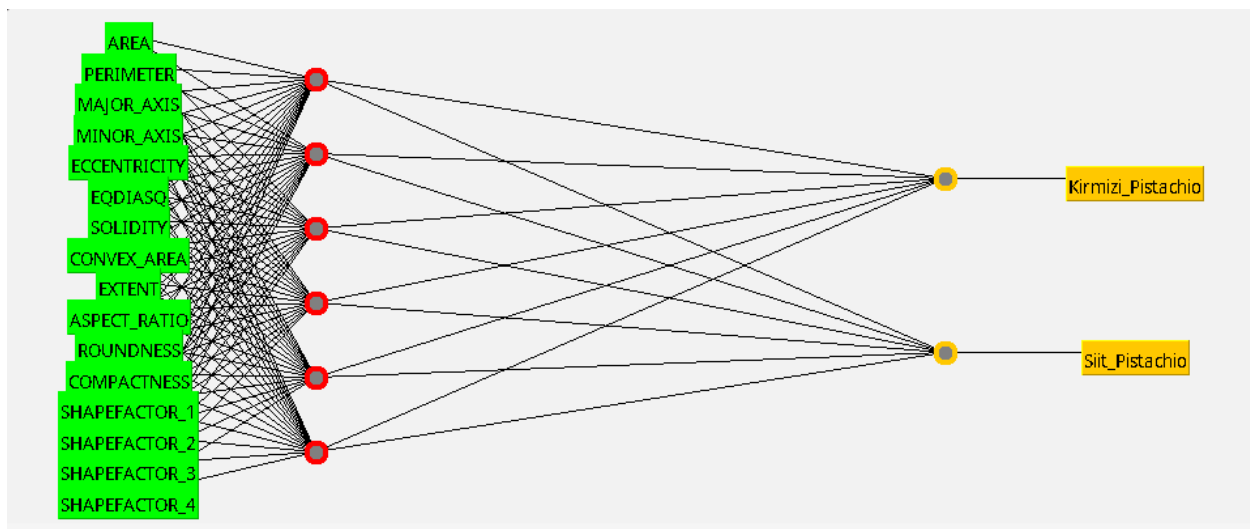
En cada celda, el primer valor corresponde al Root Mean Squared Error y el segundo al Root Relative Squared Error.

Nº Capas Ocultas	Training Set - 2000 epochs	Test Set	Test Set Config	Cross Validation	Cross Validation Config
0	0.3015 60.4394 %	0.3309 62.0622 %	0.3283 61.5753 %	0.3098 62.6492 %	0.3109 62.857 %
1	0.3114 62.4145 %	0.3363 63.0736 %	0.3409 63.9485 %	0.3187 64.4352 %	0.3181 64.3212 %
2	0.3007 60.2759 %	0.3502 65.691 %	0.3492 65.5045 %	0.3165 63.9861 %	0.3167 64.037 %
3	0.2869 57.5123 %	0.3244 60.8551 %	0.3398 63.7371 %	0.3148 63.6503 %	0.3162 63.9356 %
4	0.2856 57.2562 %	0.3102 58.1912 %	0.3394 63.6659 %	0.3112 62.9314 %	0.3074 62.1467 %
5	0.2871 57.5493 %	0.3038 56.9754 %	0.3379 63.3815 %	0.3084 62.3641 %	0.3136 63.4145 %
6	0.2809 56.3086 %	0.3131 58.7268 %	0.3371 63.2375 %	0.3128 63.2477 %	0.3137 63.4294 %

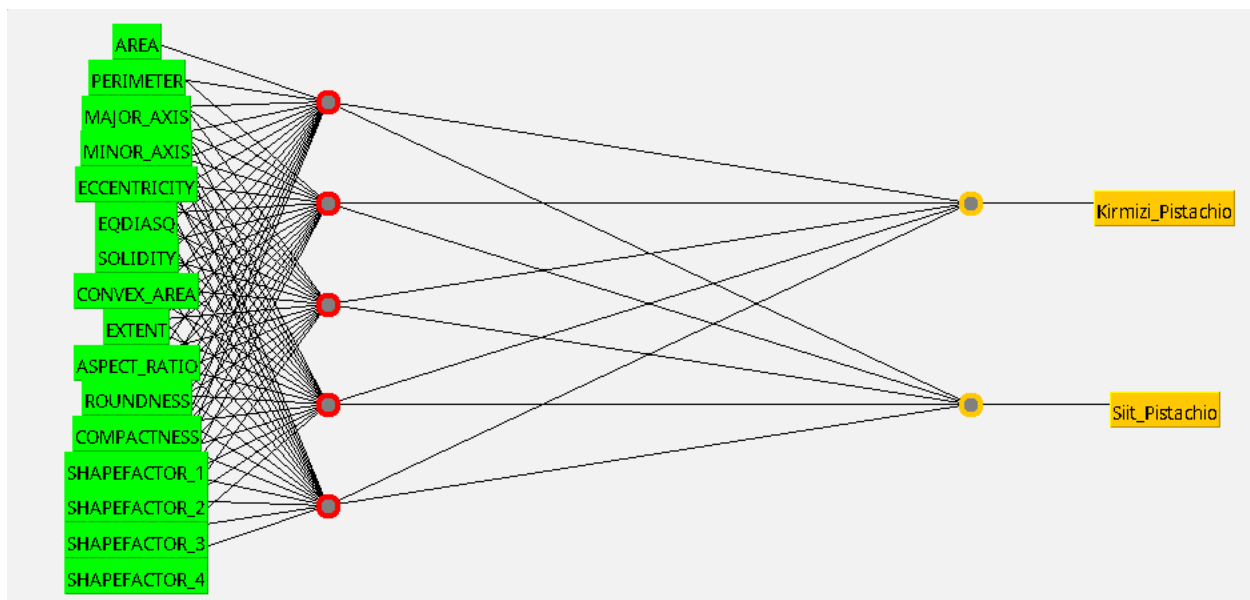
1. En el primer caso, la mejor red neuronal es la que tiene 6 neuronas ocultas ya que al aumentar el número de neuronas el error se disminuye. Por ejemplo, en este caso no se da el caso de



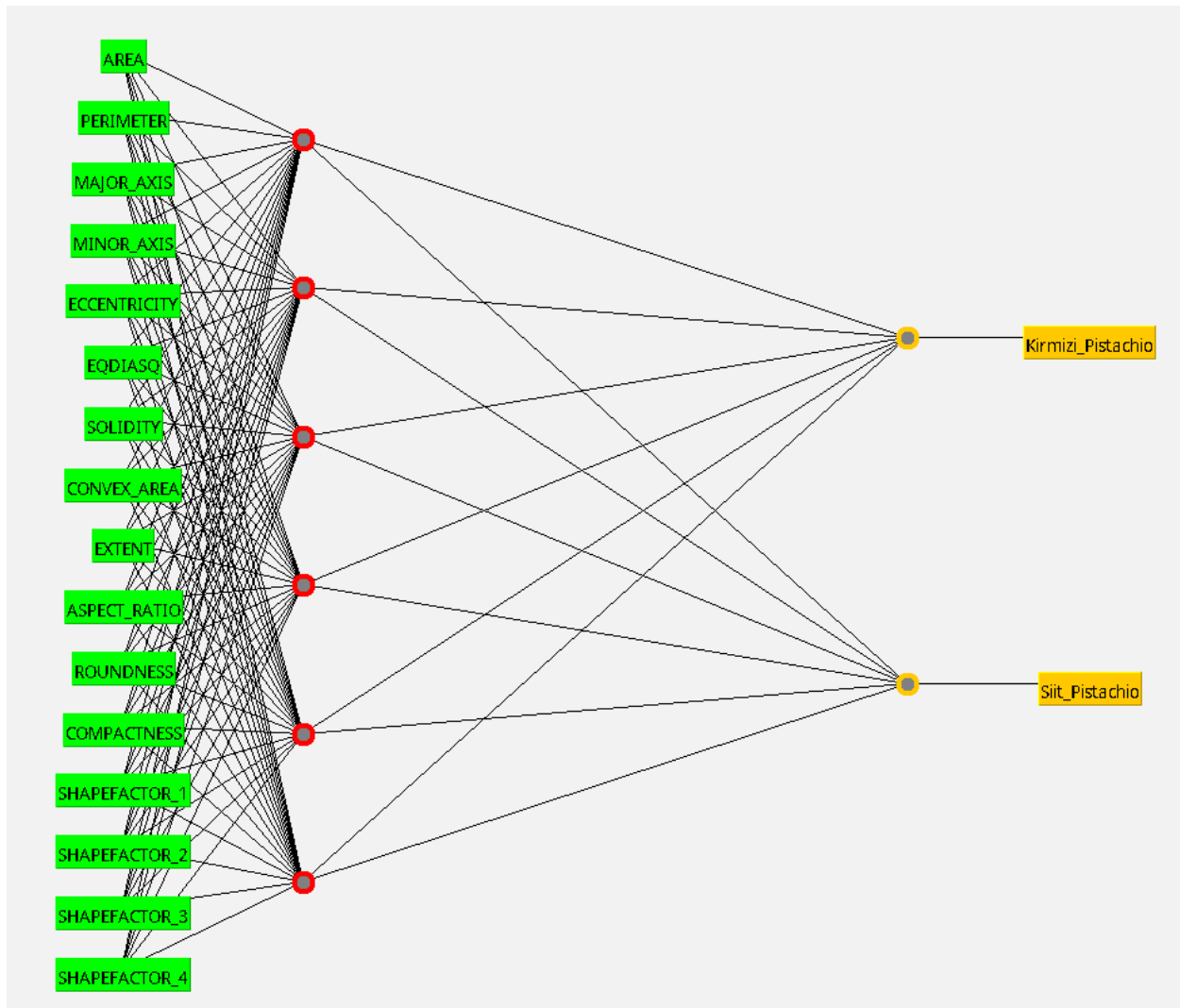
sobreajuste, pero si seguimos aumentando puede que descubramos que es un subajuste. Además, esta es la mejor red neuronal de las 5.



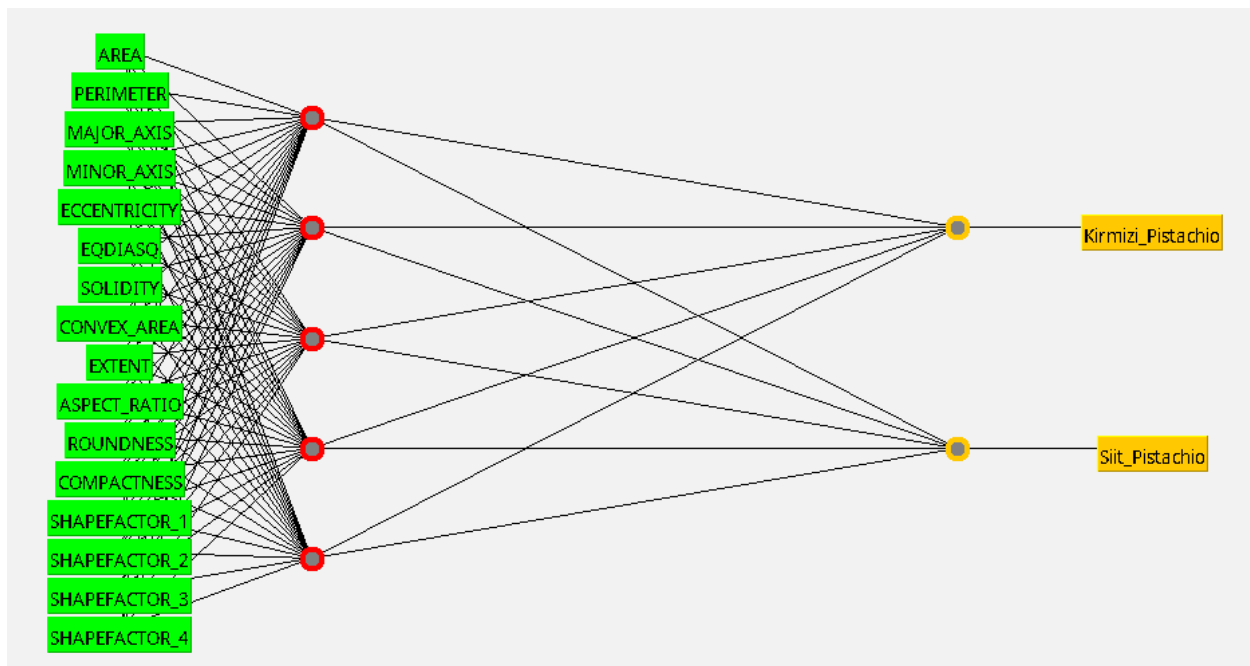
2. En el segundo caso, observamos que el mínimo error lo obtiene la red con 5 neuronas, ya que en la que tiene 6 neuronas empezamos a ver que el error aumenta, lo que quiere decir que hay un sobreajuste.



3. En el tercer caso sucede como en el primero, si seguimos aumentando puede que nos encontremos con que se trata de un subajuste.



4. En este caso se nos da la situación del segundo caso, donde observamos que cuando añadimos una nueva neurona, se produce un sobreajuste.



5. Por último, el mejor caso es la de 4 neuronas, ya que al aumentar el número de neuronas el error aumenta, detectándose así un sobreajuste.

